A SPATIAL ECONOMETRIC ANALYSIS OF CROP INSURANCE, CLIMATE CHANGE, AND US CORN ACREAGE

A Thesis

Presented to the Faculty of the Graduate School

of Cornell University

In Partial Fulfillment of the Requirements for the Degree of

Master of Science

by

Alyssa P Miller

August 2015

ABSTRACT

A spatial econometric approach is employed to investigate the impact of insurance subsidies and expected growing season weather on corn acreage in the United States. Increases in insurance subsidies and expected returns on insurance have a marginal impact on planted acreage in the Corn Belt. Expected temperature and precipitation, typically overlooked in the literature, are significant determinants of planted acreage. Furthermore, acreage response to temperature varies according to latitude and may partly explain the increases in corn acreage in the Northwestern Corn Belt over time. The spatially heterogeneous relationship between temperature and planted acreage response has important implications for acreage choices under various climate change scenarios.

BIOGRAPHICAL SKETCH

Alyssa P Miller (Alyssa Marie Pizzolanti) was born November 10, 1985 in Syracuse, New York. She graduated from Cornell University with a Bachelor of Science in Atmospheric Science and a minor in Applied Economics and Management in May 2007. After graduating, she spent five years as an equity finance trader at Barclays Capital Investment Bank. At the time of her departure she was a Vice President in the firm's Prime Services division. She is a Master of Science candidate in Applied Economics and Management at Cornell University, and has accepted a position as an Associate Director in hedge fund risk valuation and leveraged portfolio pricing at Scotiabank.

ACKNOWLEDGEMENTS

Thank you to my husband Andrew Miller, my parents James and Janice Pizzolanti, my maternal grandmother Maryann Marciniak (Granny), and my in-laws Steven and Virginia Miller. Without your support and encouragement I would not have been able to finish this Master of Science degree.

Of course none of the research I have produced whilst at Cornell would be possible were it not for the support and guidance of my advisor and committee chairman, Dr. Joshua Woodard. Words cannot express how grateful I am for your insight and your patience throughout the course of my Master program. I have learned so much from you and I look forward to continued collaboration. Furthermore, I would like to thank two members of Dr. Woodard's core research group, Dr. Leslie Verteramo Chiu and Dr. Yann de Mey. It has been a pleasure working with you and your insight and assistance has been paramount to my success whilst at Cornell.

I wish to acknowledge a few other faculty and staff members who, in addition to Dr. Woodard, had a profound impact on me whilst at Cornell and made my experience so enriching. Dr. David Just, thank you for serving as a member of my thesis committee and for your insight which has certainly ensured accuracy in the modeling of this thesis and provided credibility to the results. Dr. Calum Turvey, thank you for giving me the opportunity to travel to Nanjing, China to present an earlier version of this thesis at the first International Academic Conference for Graduate Students at Nanjing Agricultural University. Dr. Alison Morrill Chartrill, thank you so much for the opportunity to work with the Cornell Institute for Climate Change and Agriculture. Finally, I want to thank Linda Sanderson for all of her assistance throughout the past two years. Your work behind the scenes does not go unnoticed and is appreciated by everyone in the department!

My time at Cornell University was made possible in part due to the financial contributions of the department of Applied Economics and Management and the research grants of Dr. Joshua Woodard. This generosity is appreciated I am very thankful for its contributions to my education.

TABLE OF CONTENTS

CHAPTER I

INTRODUCTION

Statement of the Problem

Much of the recent agricultural economic literature has US federal policy as a central theme. Some of this literature is critical of existing policy and calls for dramatic changes to be made to program structure. For example, some economists have argued that the availability of insurance incentivizes, as part of the Federal Crop Insurance Program (FCIP), encourages agricultural production on land that otherwise might not have been planted, resulting in economic inefficiency. These authors call for either a reduction or complete elimination of premium subsidies for agricultural insurance. On the other hand, some literature demonstrates the need for federal policy in light of the free market's failure to address a particular economic problem. Such is the case for economists who model agricultural yields as a function of weather and predict dire yield outcomes in the future under various climate change scenarios if policy is not implemented to attenuate the impact mankind has on the climate.

At the intersection of these two themes, critics of the FCIP contend that the availability of subsidized crop insurance reduces the motivation for farmers to adapt to a changing climate, and is therefore a detriment to the long-term sustainability of agriculture (Wright, 2014). Conversely, proponents of the FCIP view federal support of crop insurance necessary in light of the market's failure to provide a robust insurance market for agricultural products (Kramer 1983; Miranda and Glauber, 1997; Skees and Barnett, 1999; Zacharias and Collins, 2013). These authors contend the FCIP is essential to a robust agricultural risk management program that can remain financially stable, globally competitive, and resilient if there are dramatic changes to the climate in the US Midwest in the future.

To analyze the effect of FCIP subsidies on agricultural production, many agricultural economists have modeled acreage *enrolled* in the FCIP as a function of federal policy. Less attention has been placed on modeling acreage *planted* as a function of these factors. And while there has been a great deal of attention placed on modeling agricultural yields as a function of weather, weather has been overlooked as an explanatory variable when modeling planting behavior. In fact, there lacks a comprehensive analysis of planted acreage on a large geographic and temporal scale which adequately accounts for prices, yield and yield risk, production, urbanization, numerous changes to federal policy including the FCIP, and changes in expected weather. Moreover, many of the previous acreage models literature ignore spatial relationships, in particular county land size heterogeneity and the instability of the relationship between the covariates and planted acreage over space.

In this paper we develop a spatial econometric model of planted acreage to determine the impact of the FCIP and expected growing season weather on farmer planting behavior. The crop of focus for the acreage model is corn, given the breadth and depth of production and financial data available, and because the US is the world's largest producer and exporter of corn.[1] Additionally, US corn producers are the largest consumers of federal crop insurance in terms of premium volume (RMA Summary of Business, 2014). Production is analyzed specifically in the US Midwest because the Midwest accounts for the majority of US corn production, and production methods (i.e. little acreage is irrigated) and loss experience are relatively homogenous. Midwest corn farmers typically make their planting choices, procure their seeds for planting, and make insurance elections by March, with planting beginning in late April or early May.

[1] Source: Economic Research Service. http://www.ers.usda.gov/topics/crops/corn/trade.aspx. Accessed on October 1, 2014.

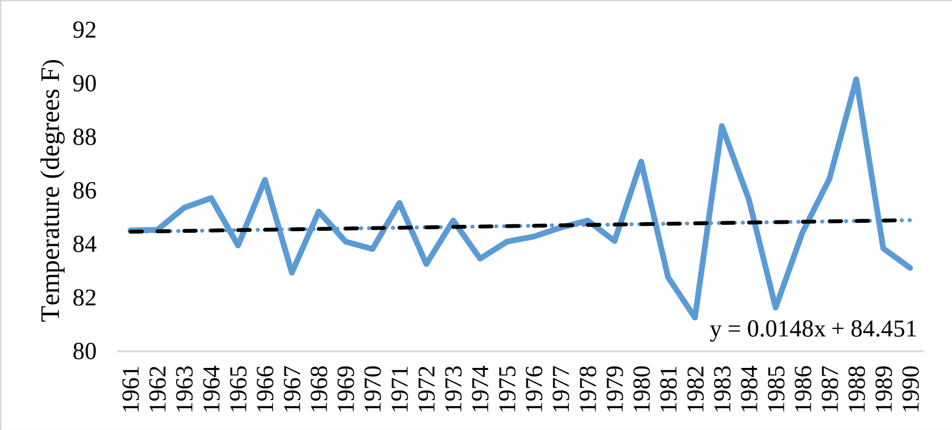
Background

Since the 1980s, the United States (US) government has made several increases to crop insurance subsidies in order to encourage greater participation by farmers and reduce or eliminate the dependency of US agriculture on direct payment, ad hoc disaster assistance, and price support programs. Because of the increases in subsidies over time, critics of the FCIP claim it is not a risk management tool but rather a disguised income payment program (Goodwin, 2001). Furthermore, with recent budgetary pressures due to national debt concerns coinciding with a period of record-high farm incomes and land values, federal expenditures on FCIP subsidies and pricing have been scrutinized. In a recent evaluation of the FCIP, Glauber (2013) calculated that for 200,000 farms, every dollar paid into the FCIP in crop insurance premiums resulted in an expected return of $1.90 from 1990 to 2011. Glauber estimated that it cost the government $1.10 for farms to be paid the $0.90 in profit. While much of the literature analyzing planted acreage decisions control for the obvious explanatory variables such as prices, yields, yield risk, and urbanization, often federal policies such as the FCIP, the Conservation Reserve Program (CRP), Farm Bills, and ethanol mandates are ignored.

Climate change has attracted a great deal of attention in recent agricultural economic literature, namely in regard to increased yield risk and future food security concerns (see e.g. Gregory, Ingram, and Brklaicich, 2005; Iglesias et al., 2007; Schlenker and Roberts, 2009; Lobell et al., 2011, among others). The focus on this literature has been to use parameter estimates recovered from econometric model to forecast yield outcomes across various climate change scenarios. Yet it is often overlooked and unacknowledged that climate change, and warming in particular, has occurred over the historical time series that these models were estimated. In the US Midwest the average air temperature increased by more than 0.5 degrees

Celsius between 1900 and 2010, with warming accelerating from 1950 to 2010.[2] For example, in McLean County, Illinois, a primary county for US Midwest corn production, there is a summertime temperature warming trend of 0.014 degree Fahrenheit per year (Figure 1).

Figure 1. Average Summertime Maximum Temperature (June, July, August) from 1961-1990 for McLean County, Illinois



Even the most conservative climate change scenarios for 2050 predict higher springtime precipitation, higher summertime temperatures, and reduced late-summer precipitation in the US Midwest (Figures 2 and 3), all potentially detrimental to field crop production. This has motivated the extensive body of literature regarding yield performance as it relates to weather. Yet literature linking farmer planting decisions to climate change is scant, and little consideration has been given regarding farmers' expectation of weather in the upcoming growing season (which may be influenced by perceived or realized climate change) and its impact on farmer planted acreage. Over a longer time series (20+ years) we posit this may be a key explanatory factor explaining trends in increased acreage in counties northwest portion of the Corn Belt, with historical climate change resulting in warming that was *beneficial* to agriculture in this region. The model developed in this thesis will serve to support or reject our hypothesis, and allow us to gain insight as to the planting response in the future if the climate continues to warm.

---

[2] Source: National Climate Assessment 2014 Regional Report for the US Midwest. http://nca2014.globalchange.gov/report/regions/midwest. Accessed on 2/2/2015.

Figure 2. Predicted Temperature Departures (degrees Fahrenheit) in 2050 from 1961-1990 Average Temperature for McLean County, Illinois[3]

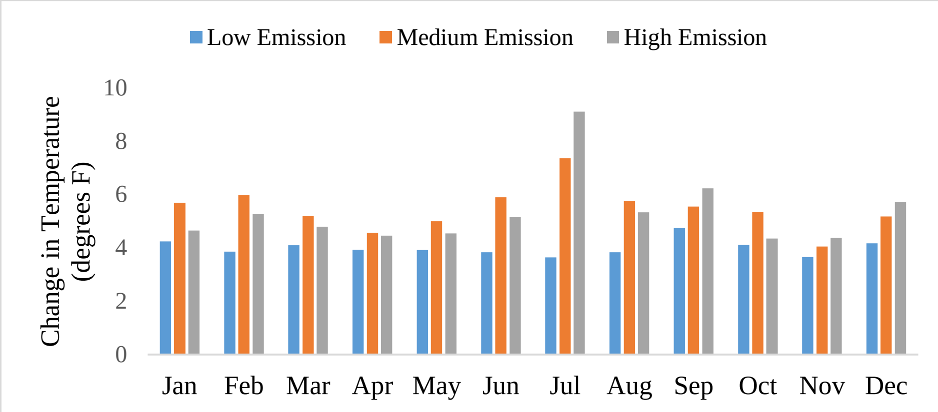
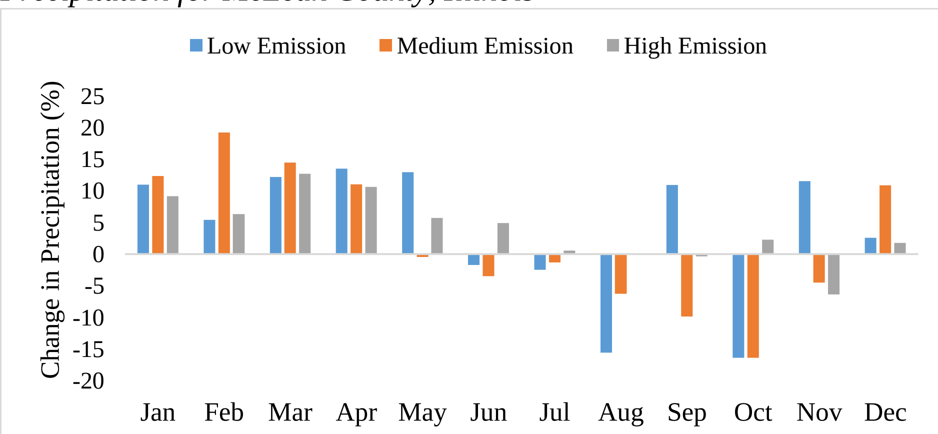


Figure 3. Predicted Precipitation Departures (%) in 2050 from 1961-1990 Total Monthly Precipitation for McLean County, Illinois



Research Objectives

We utilize spatial econometric techniques to regress county-level indexed corn acreage planted on lagged indexed corn and soybean acres, expected corn and soybean yields, corn and soybean yield risk, a regional price to cost index for corn production, a relative price ratio of corn to soybeans, urbanization, Farm Bill time dummies, crop insurance subsidies, expected loss

[3] Figures for McLean County, Illinois are generated using data from ClimateWizard.org. Climate Wizard processes climate change projections generated by 16 different general circulation models (GCMs) for the 2050s and the 2080s. The sixteen general circulation models predicting mean changes in temperature and precipitation for 2040 - 2069 and 2070 - 2099 as compared to a 1961 – 1990 base period are: BCCR-BCM2.0, CGCM3.1 (T47), CNRM-CM3, CSIRO-Mk3.0, GFDL-CM2.0, GFDL-CM2.1, GISS-ER, INM-CM3.0, IPSL-CM4, MIROC3.2 (medres), ECHO-G, ECHAM5/ MPI-OM, MRI-CGCM2.3.2, CCSM3, PCM, and UKMO-HadCM3. Graphs are generated at the 50th percentile consensus among these 16 models across three possible emission scenarios (low, medium, high).

experience for crop insurance, lagged ethanol consumption, and expected growing season temperature and precipitation, in order to investigate the key drivers of US Midwest farmers' acreage decisions.[4]

The model developed in this thesis satisfies several research objectives. First, we determine whether there is a high degree of spatial dependence or spatial heterogeneity in US Midwest corn acreage planted that warrants the use of spatial econometric methods. Second, we assess whether crop insurance availability, subsidies, and expected returns have an economically significant influence on US Midwest corn acreage. The acreage model will allow us to quantify US Midwest corn farmers' historical acreage response due to changes in expected temperature and precipitation, and use these estimates to predict acreage response under future climate change scenarios. We use a quantile regression to determine whether there is a statistically significant difference in acreage response to changes in the FCIP, US energy policy, and expected weather between US Midwest counties that consistently plant a high amount of corn acreage versus counties that transition toward or away from planting corn. Finally, the acreage model constructed in this thesis allows us to analyze the difference in acreage response by US Midwest corn farmers and corn farmers in fringe-producing states to changes in the FCIP, US energy policy, and expected weather.

Organization of Thesis

This thesis is organized as follows. In Chapter 2 the existing literature is reviewed and contributions of this research to the existing literature are identified. We discuss data sources and variable selection in Chapter 3. In Chapter 4 we discuss the econometric framework for the empirical analysis. We interpret the model results, perform robustness checks, validate the

[4] County-level corn and soybean acreage is indexed by the maximum amount of acreage planted during the time series to control for county land size and topography.

model, and perform scenario analyses to aid us in interpreting the complete model in Chapter 5. Concluding remarks and suggestions for future research are made in Chapter 6.

Summary of Findings

Lagged acreage, yield risk, the relative price ratio of soybeans to corn, the ratio of prices received to input prices paid, and lagged ethanol consumption are statistically significant explanatory factors that impact farmers' acreage decisions.

Expected loss ratios net of subsidies are statistically significant, but of marginal economic significance unless there are dramatic changes to the FCIP, such as the complete elimination of subsidies. Even under a scenario where program subsidies are eliminated and program pricing is adjusted such that the average expected loss ratio net of subsidies for the sample decreases from 1.445 to the actuarially fair rate of 1, there would be a less than 1 point change to indexed acreage in each county. Models using expected loss ratios net of subsidies instead of expected subsidies on a per bushel basis and expected gross loss ratio were more robust to out of sample validation techniques.

Expected temperature and precipitation in the growing season are statistically significant explanatory factors driving farmer acreage decisions and are jointly considered by the farmer. Furthermore, the effect of a change in temperature on acreage varies by county latitude.

The results of the quantile regressions suggest that the effect of lagged corn acreage, corn prices relative to soybeans, lagged ethanol consumption, and expected loss ratios vary according to the quantile of the dependent variable. For the lowest quantile of planted acreage the parameter estimate for lagged corn acreage is close to 1, whereas in the highest quantiles lagged acreage is less of a determinant of acreage in the upcoming season. The lowest quantiles are also the most responsive to changes in the relative price of corn to soybeans. Counties in the lowest quantile are more apt to reduce corn acreage when the market price of soybeans increases relative to corn. The effect of lagged ethanol consumption on planted acreage is positive only for the highest quantiles of planted acreage; hence, increased ethanol demand can only be attributed to increases in planted acreage in counties already planting close to 100% of their maximum historical corn acreage.

CHAPTER II

LITERATURE REVIEW

There is more than sixty years of literature by agricultural economists on the subject of planted acreage. There is a mix of econometric frameworks including recursive models, expected profit maximization, traditional rent theory, and dynamic optimization, and specifications such as OLS and Logit. Previous planting, prices, yield performance, and yield risk have been identified as the primary drivers of planted acreage. Insurance availability and subsidies have been shown to have a statistically significant but marginal impact on planted acreage. The influence of expected weather on planted acreage has for the most part been overlooked.

In spite of the extensive literature on the subject of planted acreage there remain gaps in the literature to address. By using a spatial econometric framework we account for the spatial dependence and heterogeneity present in the data, either not properly accounted for or completely ignored in the previous literature. We use a quantile regression to expand upon previous analyses which only evaluated the effect of explanatory variables on mean acreage. In particular, we analyze whether the acreage response to insurance subsidization or expected weather differs for counties that historically plant a large proportion of county acreage in corn versus those that transition toward or away from planting. We utilize a larger spatial and temporal scale than previous studies in order to analyze the influence of longer term changes in climate on planted acreage. Finally, improve upon the specifications used for insurance subsidies and weather in comparison to the previous literature.

A classical acreage model from which much of the modern literature on planted acreage has evolved is the Nerlovian model. Nerlove (1956) proposed a simple model of planted acreage as a function of expected price,

$$x_t = \alpha_0 + \alpha_1 P_t^* + \mu_t \tag{1}$$

where x_t is acres planted in time t, and P_t* is expected price. Furthermore, he posited that expected price was a function of previous year's prices. That is,

$$P_t^* = \beta P_{t-1} + (1-\beta)\beta P_{t-2} + (1-\beta)^2 \beta P_{t-3} + ... \tag{2}$$

where β is an autoregressive coefficient. By substitution, acreage in time t becomes a function of both price and acreage in the previous time period.

$$x_t = \pi_0 + \pi_1 P_{t-1} + \pi_1 x_{t-1} + v_t \tag{3}$$

$$\pi_0 = \alpha_0 \beta$$
$$\pi_1 = \alpha_1 \beta$$
$$\pi_2 = 1 - \beta$$

Nowshirvani (1971) incorporated yield risk into the Nerlovian model by contending farmers' expected prices are also shaped by expected yields. Just (1974) argued that risk, measured by the variance of revenue, should influence farmer acreage decisions. Tegene et al. (1988), in an examination of Iowa corn acreage, modified the classical Nerlovian model to consider not only lagged acreage for corn, but also lagged acreage for other crops to capture crop rotations and substitution effects.

Increases in crop insurance availability and subsidies have been attributed to changes in farmer production patterns by altering the quantity and allocation of available acreage to crops where insurance is offered and more heavily subsidized (see e.g. Wu, 1999; Young and Westcott, 2000; Goodwin and Smith, 2003). From an environmental perspective, increased subsidies were linked to farmers' choice to bring riskier and environmentally sensitive land into production (see e.g. Wu and Adams, 2001; Lubowski et. al, 2006). Yet the prevailing conclusion has been that while the impact of crop insurance availabilities and subsidies is statistically significant, it is not necessarily economically significant. Dramatic increases in crop insurance subsidies with the

1994 FCIRA and ARPA in 2001 were found to have resulted in statistically significant but economically modest conversion of land for farm use (see e.g. Vesterby et al., 2002; Goodwin, Vandeveer, and Deal, 2004; Lubowski, et al., 2006; Claassen, Cooper, and Carriazo, 2011; Walters et al., 2012). For example, using corn and soybean data for the Corn Belt states in the 1980s and 1990s, Goodwin, Vandeveer, and Deal (2004) found concluded shocks to insurance premiums for corn and soybeans of +/- 30% result in a large changes in insurance participation (measured as the ratio of liability to maximum possible liability) of +/-20%, but less than +/-1% responses in acreage. Lubowski et al. (2006) suggest the increase in crop insurance subsidies with the passage of FCIRA in 1994 changed land use measurably, but modestly; the change in premium subsidies pre and post 1994 resulted in a 2.5 million acre (0.82%) increase in cultivated cropland in the US. Miao, Feng, and Hennessey (2011) found large increases to insurance subsidies resulted in minimal impacts to acreage relative to other acreage drivers such as price.

Young, Vandeveer, and Schnepf (2001) found that acreage response due to changes in crop insurance subsidies and expected indemnities vary by crop and by region. The authors conclude regional response and crop mix adjustments dampen attribution of the overall acreage response to the FCIP. Consequently, we estimate our model for both the primary Corn Belt as well as fringe producing regions in order to compare acreage response due to the FCIP.

Westcott and Young (2004) contend coupled farm programs impact farmers' acreage decisions.[5] They argued that decoupled programs do not impact production choices since payments are fixed and not linked to production choices and/or output levels, whereas coupled programs were closely linked to a farmer's crop choice and acreage allocation. However, decoupled programs may have indirectly impacted acreage and production because the extra

[5] A farm program is *coupled* if there is a direct link between the benefit received and the farmer's production and market conditions (i.e. prices). In contrast, a *decoupled* payment would be a fixed payment similar in nature to the direct payment program.

income stream changed producer wealth, resulting in increased farm investment and changing farm tolerance to risk. This notion of changing risk profiles is echoed by Wu and Adams (2001) who used an expected utility maximizing framework to conclude that revenue insurance programs reduce farmers' production risk by guaranteeing a revenue floor. This results in a censored distribution of crop revenue, thereby increasing the expected value of revenue and decreasing the variance of revenue. Thus insurance, particularly revenue insurance which provides both price and yield protection, should impact acreage decisions as it changes the risk profile of a crop. Subsequently in this analysis we include a time dummy to coincide with the introduction of revenue insurance for corn as well as other key federal Bills and Acts which increased program subsidization.

Goodwin, Vandeveer and Deal (2004) analyzed the impact of the CRP and FCIP in the context of three choices an agricultural producer faces: (1) what to produce, (2) how to produce, and (3) whether or not to participate in an available government program. The authors used acreage planted as the dependent variable in their model of choice (1). Acreage planted regressed on county size, rather than some measure of indexed acreage that accounts for county size heterogeneity. The authors' measure implicitly imposes the same marginal effect of county land size on each county in the sample. A measure of insurance participation is included in their acreage model, equal to liabilities in a county observed/maximum possible liability. The maximum possible liability in each county is defined as a ten year rolling historical average of price*acreage*75% yield. This variable may confound the effects of risk aversion and subsidization. A model that explicitly includes subsidies and loss ratios is preferred from a policy perspective since subsidies are the most contentious component of the program.

Lubowski et al. (2006) used the 1994 FCIRA is used as a natural experiment to observe how land-use conversions change in response to crop insurance subsidies. They use 1992 to 1997 as s time series in order to identify the impact of this policy change by comparing land-use changes before and after the passage of the 1994 FCIRA. The model builds on traditional rent theory to estimate the likelihood that a parcel of land in a particular land use in 1992 either remains in the same use or is converted to a different use by 1997. Lubowski, Plantinga, and Stavins (2008) adopted a dynamic optimization approach where the landowner is assumed to make land use decisions according to that which maximizes the present value of stream of net expected benefits of the land. For simplicity it was further assumed land use is a fixed and irreversible decision. An infinite horizon dynamic optimization framework with an irreversible decision for land use may not be the best specification; although it is often expensive and timely to transition land, shifts toward and away from farming has been observed in the Midwest.

Huang and Khanna (2010) analyzed crop yields and crop acreage using county-level data from 1997-2007. The focus of the analysis was to estimate yields and acreage while properly accounting for climate variability, technology, and crop prices. Adopting a Nerlovian approach, the authors define acreage as a function of lagged acreage, price, price risk, yield, yield risk, population density, a linear time trend, and a climate trend term. This model does not account for critical federal policies that impacted acreage over the time series; namely crop insurance and energy policy. Furthermore, the authors do not account for vast heterogeneity in county size.

Chen and Onal (2012) posit acreage elasticities are inelastic on intuitive grounds, as supply of input factors for agricultural production (land and water) are inelastic, demand for food and fiber is inelastic, and because soil quality, climate conditions, and technology factors lead to inflexible production practices and land use decisions. The authors contend farmers' risk aversion to status quo planting behavior since farmers prefer to plant with which they are experienced adopting new alternatives. They used a double log model to estimate acreage elasticity for US corn, soybeans, and wheat, regressing county level acreage on lagged acreage, lagged commodity prices, commodity stock levels, input prices (fuel and fertilizer), and weather. The authors modeled two periods, 1977 to 1995 and 1996 to 2007. In the first of these two periods, the US farm programs strongly affected producers' acreage decisions, while acreage responses in the latter period were determined primarily by market forces such as demand for ethanol. This analysis highlights the necessity of accounting for federal policies, even over short time series.

Interestingly, there is no consensus in the literature regarding the appropriate spatial or temporal aggregation of weather data when included in a model of either planted acreage or production. The existing literature relies on observed weather in models of agricultural production. Although this approach is appropriate for yield models, it is not appropriate when modeling planted acreage. Because of inter and intra-annual variation in weather, observed data may not be a good proxy for farmers' expectations of weather. Kaiser et al. (1993) used a revenue maximization framework to posit that farmers choose cultivars (and thus indirectly make their planting decisions) according to expected yield performance, which is a namely a function of weather. While the authors concluded that climate change effects yield risk and hence farmers' expected revenue, they do not explicitly link this back to aggregate changes to corn acreage. Huang and Khanna (2010) use observed weather in their acreage model. However, the model suffers from misspecification as the observed weather incorporated in their model takes place after the planting decision has been made. The authors' climate variables are summertime temperature, precipitation, and growing degree day accumulation in year *t.* As planting decisions are made in late winter and early spring, a more appropriate specification would have been to use lagged weather. Chen and Onal (2012) include weather in their acreage model, yet there is no explicit definition of the weather variables used, and among their two time periods, 1977 to 1995 and 1996 to 2007, the later time period may be too short to father much insight regarding climate change as a driver of changes in acreage.

CHAPTER III

DATA

Data Sources and Sample Selection

Data for the model is obtained from the Agricultural Database ("AgDB") at Cornell University, a comprehensive open-source database containing economic, financial, and environmental data.[6] County-level acreage and production data are sourced from the National Agricultural Statistics Service (NASS). Historical Chicago Board of Trade (CBOT) futures prices for corn and soybeans are sourced from Quandl.[7] Regional gross production and total cash cost indicators are sourced from the Economic Research Service (ERS) Annual Commodity Costs and Returns report. County-level insurance data are sourced from RMA's Summary of Business report. State-level urbanization indicators are obtained from the US Census Bureau. Soil composition and quality information is reported by the USDA National Resource Conservation Science (NRCS). Ethanol data are obtained from the US Energy Information Administration.

The twelve contiguous "Corn Belt" states are employed in the base analysis (Illinois, Indiana, Iowa, Kansas, Michigan, Minnesota, Missouri, Nebraska, North Dakota, Ohio, South Dakota, and Wisconsin) from 1985 to 2013. The 12 states chosen for this analysis have a total of 1,051 counties for a total of 30,479 possible observations across 29 years. We also re-estimate the model for the less productive states of Colorado, Kentucky, Mississippi, North Carolina, New York, Pennsylvania, Tennessee, and Texas.

Dependent Variable Selection

[6] AgDB was created by Dr. Joshua D. Woodard and researchers at Cornell University. https://agfinance.dyson.cornell.edu/AgRiskManagement/About.
[7] Quandl is an open-source web platform containing 10 million financial datasets from 500 sources. https://www.quandl.com/.

The dependent variable for the regression is the number of corn acres planted in a county i = 1…N in time t = 1…T, indexed by the maximum number of acres of corn planted in the county over the time series. This results in a fractional explanatory variable which is bounded from 0 to 1. We then multiply this percentage by 100 for ease of interpretation of estimated regression coefficients, and our bounds become 0 and 100. That is,

$$ACRESCORN_{it} = \frac{AcresPlanted_{it}}{\max(AcresPlanted_i)} * 100 \tag{4}$$

This is in contrast to the existing literature which either ignores county land size heterogeneity (see e.g. Huang and Khana, 2010; Chen and Onal, 2012), or defines planted acreage as a function of county land size, thereby imposing at the margin the same effect of land size on planted acreage (see e.g. Goodwin, Vandeveer, and Deal, 2004). Not only does our dependent variable relax this assumption, but our variable better controls for unobserved factors that impact the viability of county acreage for farm production, such as soil quality, county-level urbanization, and topography.

Explanatory Variable Selection

Lagged Corn Acreage

We include the dependent variable lagged by one year in the regression.

Lagged Soybean Acreage

Crop rotation is important for corn because producing corn following corn on the same unit of land in the prior year slowly depletes soil organic matter and increases corn-plant pest control problems, so that expected com yield per unit of land declines. When corn follows soybeans or other leguminous plants, corn yield per unit of land is higher than with continuous corn because legumes fix nitrogen and improve soil drainage. Typically most crop rotation with corn is done with soybeans in this region. County-level normalized planted acres of soybeans lagged by one year are included in the regression. Similar to our dependent variable, this term will be left and right censored at 0 and 100 respectively. We hypothesize the coefficient for lagged soybean acreage will be positive, as more soybeans planted in the previous period should result more corn planted in the current period, reflecting crop rotation practices.

Conservation Reserve Program (CRP) Acreage

The CRP pays a yearly rental payment on voluntarily enrolled acres in exchange for farmers removing environmentally sensitive land from agricultural production and planting species that will improve environmental quality. That is, acreage dedicated to the CRP program cannot be used to plant corn or soybeans. We include CRP enrolled acres normalized by county acres in each county/year in our analysis to control for rental payments on land that may have incentivized farmers to reduce farm acreage dedicated to corn. Similar to our dependent variable, we multiply the fraction of acres in the county dedicated to CRP by 100 for ease of interpretation.

We choose this variable for our model in lieu of other government price support payments, CRP payments, wetland conservation, and other voluntary program payment data because these payments are only available in census years. Since these payments are often counter-cyclical, standard linear or cubic interpolation between years may not be appropriate. Instead we use CRP acreage as a proxy for voluntary and price support payments since enrollment figures are available annually and this variable will be highly correlated to non-insurance subsidy payments made to farmers.

Expected Corn Yields

County-level expected corn yields (calculated as production in bushels divided by the number of acres planted) are included in the analysis. To construct expected yields we fit a simple linear trend to observed historical yields from 1975 to time *t*-1. Using the parameters recovered from this simple regression we calculate the fitted yield for time *t* and use this as a proxy for farmer-expected yields. By using this measure instead of the average of previous historical yields we not only account for county-level yield trends due to technological advancements, but we reduce the influence of idiosyncratic yields from previous years. We hypothesize the parameter estimate recovered by the regression will be positive as farmers will respond to favorable expected yield outcomes by dedicating more acreage to corn in the upcoming growing season.

Corn Yield Risk

We include county-level corn yield risk in the model, calculated as the coefficient of variation, or yield variability normalized by mean yields over the previous five years. We suspect increasing variance would result in reduced acreage planted due to farmer risk aversion.

Expected Soybean Yields

County-level expected soybean yields (calculated as production in bushels divided by the number of acres planted) are included in the analysis. These yields are constructed in the same manner as corn yields. We expect the parameter estimate to be negative, as corn and soybeans are substitutes, and a decrease in expected yield performance in soybeans may incentivize a farmer to plant additional acres of corn.

Soybean Yield Risk

We include county-level soybean yield risk in the model, calculated as the coefficient of variation, or yield variability, of soybean yields over the previous five years. The time period used for our soybean yield risk term matches that of lagged yield term, since we hypothesize farmers consider yield performance and yield risk over the same time period when making acreage decisions.

Price to Cost Ratio

In order to avoid endogeneity issues that could arise from incorporating price directly into our acreage model, we construct a ratio of regional total gross value of corn production to total cash costs. Gross production is defined as regional yield multiplied by harvest price multiplied by regional acres planted, and this figure excludes government payments. Total cash costs include seeds, fertilizer, labor, taxes, and debt expenses. For years with this ratio greater than 1, this indicates a profitable experience for the farmer. For the model, we lag this variable by one year and assume the farmer considers the previous year's outcome when making acreage decisions for the upcoming year.

Relative Price Ratio

A relative price ratio of soybeans to corn is constructed using the Risk Management Agency (RMA) price discovery period prices[8]. Since corn and soybeans are substitute crops, we suspect if soybeans are relatively rich during the price discovery period this may result in fewer acres of corn planted.

Urbanization

We use a state-level urbanization indicator produced by the US Census Bureau in the analysis. The figure represents the amount of urban population as a percentage of the overall population in the state at the time of the census. This census is taken every ten years, so there are

[8] The RMA price discovery period prices are used in pricing insurance rates for the upcoming growing season. For soybeans, the RMA discovery period price is calculated by taking a simple average of futures market closing prices from the month of February, for contracts with January expiry of the following year. For corn, the price is calculated by taking a simple average of futures market closing prices from the month of February for contracts with December expiry of the same year.

five measurements available at the state-level for our chosen time series. We apply linear interpolation for years in between each census. Values range from 0 to 100%. We hypothesize with an increasing competition for land by urban expansion there could be locations that experience declines in acreage.

Expected Weather

Given these historical and forecasted changes in temperature and precipitation in the US Midwest, it would be prudent to include some measure of farmers' expectation for weather into an econometric model of planted acreage. These values are meant to reflect farmers' rational expectation of future weather given previous experience, thereby accounting for climate change observed over the time series.

We include a measure of the farmers' expectation for weather in the growing season. The proxy used for farmer-expected temperature is a simple average of the monthly *expected* mean temperature for June, July, and August. For precipitation we include the sum of the *expected* total precipitation for June, July, and August. A quadratic temperature term is added since the effects of temperature on planted acreage has been shown to be nonlinear (Schlenker and Roberts, 2009). Temperature and precipitation are interacted since the effect of temperature on planted acreage will be dependent on precipitation, and vice-versa. Expected mean temperatures and total monthly precipitation are constructed via a spatial autoregressive model of historical monthly temperature and precipitation with agricultural district fixed effects and time trends.

We use a spatial autoregressive method to garner implied climate trends, or the *expected* weather for every agricultural district in the contiguous US. We hypothesize this measure will have reduced variability and reflect a more rational expectation of weather, and thus provide better insight and use in a model of planted acreage. Due to the high degree of geospatial correlation in weather, a spatial econometric specification is appropriate.

We utilize a spatial autoregressive model, where the temperature in agricultural district *i* in month *m* = 1,…,12, and time *t* = 1961,…,1990 is a function of a spatial lag term, ρ, the previous month's mean temperature, the previous months' total precipitation, time fixed effects, and agricultural district individual effects. The weather data for this analysis is the 4km resolution monthly mean temperature and total monthly precipitation as reported by the PRISMs Climate Group at the University of Oregon. Our time series matches the time series used to make climate predictions for the 2050s and 2080s. The analytical equation for the model is:

$$WX_{i,m,t} = f(WX_{i,m-1,t}, Time_t, \alpha_i) \tag{5}$$

Where the weather for agricultural district *i* in month *m* in time *t* is a function of the previous month's weather, time fixed effects, and agricultural district fixed effects. The formal spatial econometric specification is:

$$WX_{i,m,t} = \tau WX_{i,m-1,t} + \gamma WX_{i,m-2,t} + \rho W_N WX_{i,m,t} + Time_{i,m} + \alpha_i + \varepsilon_{it} \tag{6}$$

Where weather in district *i*, month *m*, and year *t* is a function of an autoregressive term, a spatial lag term, a matrix of time fixed effects, and a district level time invariant fixed effect. Model results for monthly mean temperature and monthly total precipitation for the planting and growing season are presented in Tables 1 and 2 below.

Table 1. Regression Results, Expected Mean Temperature

	m-1	*m-2*	*W*Y*	*loglik*	*adj. r^2*	σ^2
April	0.0239***	0.0110***	0.9859***	-5910.33	0.9959	0.1100
May	0.0235***	0.0162***	0.9879***	-5215.84	0.9945	0.0980
June	0.0213***	0.0108***	0.9929***	-4703.42	0.9934	0.0910
July	0.0309***	0.0079***	0.9869***	-5057.56	0.9901	0.0970
August	0.0344***	0.0113***	0.9779***	-4908.78	0.9915	0.0950
September	0.0323***	0.0097***	0.9859***	-5162.53	0.9932	0.0980

*Note: ***, **, * denotes statistical significance at α= 0.01, 0.05, and 0.10 respectively.*

Table 2. Regression Results, Expected Total Precipitation

	m-1	*m-2*	*W*Y*	*loglik*	*adj.* r^2	σ^2
April	0.0149***	0.0044	0.9669***	-52307.26	0.9084	242.576
May	0.0168***	0.005	0.9549***	-53330.94	0.9000	291.540
June	0.0129***	-0.003	0.9429***	-53992.51	0.8891	329.889
July	0.0071*	0.0051	0.9349***	-53722.26	0.8909	315.167
August	0.0007	0.0025	0.9369***	-53677.29	0.8808	312.384
September	-0.0055	0.0171***	0.9489***	-54655.78	0.8840	365.221

*Note: ***, **, * denotes statistical significance at α= 0.01, 0.05, and 0.10 respectively.*

In general, the model appears to have more explanatory power for temperature than for precipitation. This may reflect measurement error issues in historical precipitation data, since precipitation events often occur on a smaller scale than the resolution at which data is available. Variable coefficients for temperature are all positive and significant. For precipitation all statistically significant coefficients are positive. With the exception of September, only the previous month's precipitation is statistically significant. Spatial ρ's exceed 0.93 in all cases, indicating that even with 4km resolution there is strong spatial correlation with historical precipitation events.

Latitudinal Coordinates

Expected temperature is interacted with latitudinal coordinates under the spatial expansion specification to test our hypothesis that climate change over the time series in more northern counties resulted in increased planting, as historical warming was favorable for crop growth and development.

Insurance Availability, Subsidization, and Expected Returns

Whereas models in the previous literature have used either measures of insurance participation, such as observed liabilities over maximum possible liabilities (Goodwin, Vandeveer, and Deal, 2004), or measures of returns on insurance (Lubowski et al., 2006), our model is more comprehensive in that it incorporates all components of the FCIP that could explain changes in planted acreage, including insurance availability, subsidization, and expected returns.

Farm Bill dummies time dummies are included as a proxy for changes in insurance availability and to account for the introduction of group and revenue insurance products in 1997. In order to account for the effect of insurance subsidization and expected returns on insurance, we estimate the model using two sets of explanatory variables. First, we estimate the model using an expected subsidy-adjusted loss ratio. The loss ratio serves as a proxy for how expensive insurance is perceived to be by farmers. The expected loss ratio is calculated by dividing total expected indemnities by the subsidy-adjusted premium paid. A loss ratio greater than 1 indicates the farmer has received more payouts from the insurance product than he or she has paid in net premiums. To compute a farmer's *expected* loss ratio we must make an assumption regarding the number of years a farmer considers previous outcomes in making their prediction as to the relative costliness of insurance in the upcoming growing season. We use the five-year average historical loss ratio as a proxy for the farmer-expected expected loss ratio, since classical economic literature has shown that averages of past experience are most reflective of future experience in the case of profit-maximizing producers (Muth, 1961). We hypothesize that time periods and counties where expected loss ratios were higher than 1 may coincide with planting that is higher than the county's average acreage over the time series. Since spatial models require a balanced panel we assume the expected loss ratio is equal to the actuarially fair rate of 1 for years where RMA SOB data is unavailable (prior to 1989). For subsequent years we utilize the average loss ratio method explained above.

There are two limitations from using the subsidy-adjusted expected loss ratio as the primary measure of the effect the FCIP has on planted acreage. First, this variable only explains the joint effect of subsidies and expected returns from insurance on planted acreage. Second, this subsidy-adjusted expected loss ratio does not account for productivity heterogeneity between high yielding and fringe counties. A better specification may include subsidies and loss ratios separately and account for subsidies on a per bushel basis, rather than a per acre basis.

As an alternative measure of the impact of the FCIP on planted acreage, we include two regressors in the model; expected gross loss ratio and the *expected subsidy take*. The expected gross loss ratio is calculated as the average of the five previous years' corn indemnities divided by corn insurance premiums. We define expected subsidy take as follows:

$$ESUBTAKE_{it} = \frac{SUB_{it}}{IA_{it} * E[PX_t]} * \frac{1}{E[Yield_{it}]} \tag{7}$$

Where SUB_{it} are total corn subsidies in county *i* in time *t*, IA_{it} is the number of acres insured in county *i* and time *t*, $E[PX_t]$ is the farmer-expected price, equal to the RMA Projected Price Discovery Period price in time *t*, and $\frac{1}{\mathrm{E}[Yield_{it}]}$ is the inverse of the farmer-expected yield (production in bushels divided by planted acres) in county *i* and time t. Decomposing $ESUBTAKE_{it}$ into two parts, $\frac{SUB_{it}}{IA_{it} * E[PX_t]}$ is the price-adjusted subsidy per insured acre, which is multiplied by inverse expected yields, in order for subsidies to be expressed on a per bushel basis. This is preferred to a measure of subsidies on a per acre basis since fringe and productive regions will have different yield outcomes. Since RMA SOB data is not available until 1989, we assume subsidies, and therefore subsidy take, is equal to zero prior to 1989.

Soil

As a proxy for soil quality in this analysis, we use the National Commodity Crop Productivity Index (NCCPI) for Corn and Soybeans. Index values range from 1 (low productivity) to 100 (high productivity). The index is reflective of soil productivity for non-irrigated commodity crops only. The index relates soil content and quality, landscape, and climate factors to model the response of commodity crops.[9]

Ethanol

We hypothesize the impacts of federal energy policies enacted in 2005 (The Energy Policy Act) and 2007 (The Energy Independence and Security Act) that mandated increased ethanol usage in US automobile fuel may have incentivized greater corn planting later in our time series. We include lagged US consumption of ethanol in hundreds of millions of gallons in our analysis.

Dummy Variables

We include time dummies for Farm Bill years which may have incentivized farmers to plant more corn due to either increased insurance availability or the introduction of new insurance products, such as revenue and group insurance.

Table 3 provides a summary of all the variables used in the empirical analysis, and Table 4 presents basic summary statistics for the data.

Table 3. Regression Variables

Variable	Description	Source
ACRESCORN	The dependent variable for the analysis, county-level normalized acres planted of corn.	National Agricultural Statistic Service (NASS)
LAGCORN	County-level normalized acres planted of corn lagged by one year.	National Agricultural Statistic Service (NASS)
LAGSOY	County-level normalized acres planted of soybeans lagged by one year.	National Agricultural Statistic Service (NASS)
CRP	County-level normalized acres enrolled in the Conservation Reserve Program.	United States Department of Agriculture Farm Service Agency (USDA FSA)
EYIELDCORN	County-level average de-trended yield over previous five years.	National Agricultural Statistic Service (NASS)
EYIELDSOY	County-level average de-trended yield over previous five years.	National Agricultural Statistic Service (NASS)

[9] Additional information regarding the construction of the NCCPI Corn and Soybeans index can be found at: http://www.nrcs.usda.gov/wps/PA_NRCSConsumption/download?cid=nrcs142p2_050734&ext=pdf.

Table 3 (Continued)

CVCORN	County-level coefficient of variation of yields over previous five years.	National Agricultural Statistic Service (NASS)
CVSOY	County-level coefficient of variation of de-trended yields over previous five years.	National Agricultural Statistic Service (NASS)
PXRATIO	Relative price discovery period ratio of soybeans to corn.	Chicago Board of Trade (CBOT)
PCRATIO	The ratio of Midwest corn total gross value of production to total cash costs.	Economic Research Service (ERS)
URBAN	The percentage of urban population in a state as part of the overall population of a state.	United States Census Bureau
ETHANOL	Lagged US ethanol consumption in hundreds of millions of gallons.	US Energy Information Administration
ESUBTAKE	County-level expected subsidy take ($/bushel).	Risk Management Agency (RMA) and National Agricultural Statistic Service (NASS)
EGLR	Expected gross loss ratio. 1 for years prior to 1989, and average historical gross loss ratios for periods after 1989.	Risk Management Agency (RMA)
ELR	Expected loss ratio adjusted for subsidies. 1 for years prior to 1989, and average historical subsidy-adjusted loss ratios for periods after 1989.	Risk Management Agency (RMA)
ETEMP	County-level average expected growing season mean temperature (June-August) in degrees C.	PRISMs Climate Group
EPREC	County-level total expected growing season precipitation (June-August) in mm.	PRISMs Climate Group
SOIL	NCCPI Soil quality index for corn and soybeans.	USDA National Resource Conservation Science (NRCS)
LAT	County-level latitudinal coordinates.	United States Census Bureau
D1985	Dummy variable equal to 1 for years > 1985 and <=1990. This variable represents the 1985 Farm Bill.	
D1990	Dummy variable equal to 1 for years > 1990 and <= 1994. This variable represents the 1990 Farm Bill.	
D1994	Dummy variable equal to 1 for years > 1994 and <= 1996. This variable represents the Federal Crop Insurance Reform Act.	
D1996	Dummy variable equal to 1 for years > 1996 and <= 2000. This variable represents the 1996 Farm Bill.	
D2000	Dummy variable equal to 1 for years > 2000 and <= 2002. This variable represents the 2000 Agricultural Risk Protection Act, an amendment of the Federal Crop Insurance Reform Act.	
D2002	Dummy variable equal to 1 for years > 2002 and <= 2008. This variable represents the 2002 Farm Bill.	
D2008	Dummy variable equal to 1 for years > 2008. This variable represents the 2008 Farm Bill.	

Table 4. Summary Statistics for Regression Variables

Variable	*Mean*	*Std Dev*	*Min*	*Max*	*25th Percentile*	*75th Percentile*
ACRESCORN	67.501	22.490	1.428	100.000	54.905	83.999
LAGCORN	66.972	22.661	1.428	100.000	54.167	83.708
LAGSOY	57.339	34.808	0.000	100.000	25.369	86.725
CRP	3.007	4.176	0.000	23.357	0.251	4.029
EYIELDCORN	106.383	38.674	12.219	185.641	83.588	135.599
EYIELDSOY	34.446	11.968	0.000	58.352	29.727	41.945
CVCORN	0.223	0.179	0.000	2.236	0.113	0.276
CVSOY	0.168	0.107	0.000	1.732	0.098	0.215
PXRATIO	2.358	0.228	1.938	3.001	2.219	2.444
PCRATIO	1.133	0.277	0.637	1.647	0.944	1.321
URBAN	70.136	8.648	48.350	88.710	67.460	74.590
ETHANOL	35.726	41.945	5.103	139.291	8.664	39.044
ESUBTAKE	0.015	0.051	0.000	0.529	0.000	0.090
EGLR	0.917	0.739	0.000	4.023	0.391	1.045
ELR	1.702	1.511	0.000	8.003	0.823	2.191
ETEMP	22.128	2.196	14.699	27.595	20.606	23.785
EPREC	288.056	48.819	127.939	406.914	256.786	324.541
SOIL	47.279	19.850	5.230	91.204	31.673	62.884

CHAPTER IV

METHODS AND MODELS

OLS, Tobit, and Fixed Effect Methods

An Ordinary Least Squares (OLS) model with fixed effects can serve as a base model for comparison in a spatial analysis. However, the dependent variable for this analysis is bounded by 0 and 100. OLS may not be the best specification since predicted values from OLS are not guaranteed to generate estimated values of the dependent variable that fall within these bounds. Instead, a Type I Tobit specification with lower and upper bounds is used. OLS and Type 1 Tobit results are compared to analyze whether OLS results are significantly biased.

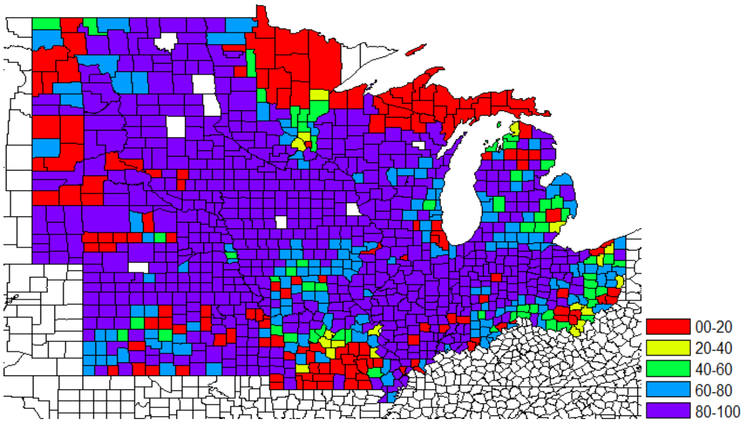
Spatial Dependence

Spatial dependence, or spatial autocorrelation, is the lack of independence between observations in a cross-sectional setting. In the case of a panel data set, it is important to note that issues of spatial dependence are not the same as issues of time dependence or serial correlation that is possibly present in the dataset.

There are two primary causes of spatial dependence in agricultural economics. First, agricultural risks faced by producers from adverse weather, pests, and diseases are all highly geospatially correlated (see e.g. Anselin, 2001; Woodard and Garcia, 2008; Woodard, Verteramo Chui, and Miller; 2014). Second, spatial dependence in agricultural economics arises from data measurement problems, such as arbitrary county boundaries.

A map of observed values of the dependent variable in 2012 is presented in Figure 4, and it appears there is spatial dependence evident in the data. A statistical test such as a Moran I Test can be used to confirm our theoretical justification for use of a spatial model in the analysis.

Figure 4. Map of ACRESCORN in 2013



Spatial Heterogeneity

In contrast to spatial dependence, spatial heterogeneity reflects a lack of stability over space of the relationship between the explained and explanatory variables of the study. That is, functional forms are not homogenous and vary with location. Often, this heterogeneity in geographical data sets can be expressed by Cartesian coordinates in space. For example, in this analysis we hypothesize the acreage response due to changes in climate (i.e. warming temperatures) will vary according to county latitude.

Spatial Econometric Methods

If there is spatial dependency and/or heterogeneity present in the data, Anselin (1988) has shown OLS estimations are no longer unbiased and consistent. As such, more sophisticated agricultural economic literature utilizes spatial econometric techniques when modeling agricultural acreage and/or yields. The advantage of using spatial econometric techniques is that a spatial model relaxes the assumption regarding independence of the error term between the counties or states to which fixed effects are applied. If fixed effects serve as a proxy for missing or unobserved data, for example, changes in soil quality over time, or improvements in seed technology, then model error terms are likely geospatially correlated.

In order to capture spatial dependence and heterogeneity in a spatial model, distance or location must be quantified. Location may be measured using either latitude or longitudinal coordinates or a distance vector, as in the case of a spatial expansion model. Alternatively, spatial model specifications may utilize a spatial weighting matrix to account for contiguity, or the relative position in space of one observational unit (i.e. county, agricultural district, or state) to another unit. Spatial contiguity can be defined in several ways, including linear, rook, bishop, and queen. We impose queen contiguity for this analysis. That is, locations are contiguous if they share either a border or a vertex.

The base notation used for a spatial weighting matrix is W_N, an N x N matrix where N is equal to the number of cross-sectional observations in the dataset (i.e. states, agricultural districts, or counties). Elements of this matrix represent the correlation between corresponding locations. For example, the value of $w_{i,j}$ represents the spatial correlation between location i and location j. Values are binary, where a 1 indicates location i and j are contiguous, and 0 otherwise. This implies by design that $w_{i,i}$ equals 1 and the diagonal of the entire weighting matrix is equal to 1. The weighting matrix may be standardized so that elements along each row of the matrix sum to 1. If location i has a total of m neighbors, then $w_{i,j}$ is equal to $\frac{1}{m}$ when i and j are contiguous, and 0 otherwise. For a large sample of counties, agricultural districts, or states, the matrix is likely to be a sparse matrix, meaning many of the elements of the matrix contain a 0. In the case of a panel data set, the spatial weighting matrix has dimensions $NT \times NT$.

The specification for a spatial lag model is:

$$\begin{aligned} &y = \rho(I_T \otimes W_N) y + X\beta + \varepsilon \\ &\varepsilon \sim N(0, \sigma^2 I_{NT}) \end{aligned} \tag{8}$$

Where y is a $NT \times 1$ vector of normalized acres planted, W_N is a $N \times N$ standardized spatial weighting matrix, I_T is an identity matrix with dimension T, and $I_T \otimes W_N$ is a $NT \times NT$ block matrix such that $I_T \otimes W_N = \begin{pmatrix} W_N & & 0 \\ & \ddots & \\ 0 & & W_N \end{pmatrix}_{NT\mathrm{x}NT}$. Here $\rho(I_T \otimes W_N)y$ represents the spatial lag term, or the influence of neighboring locations on the normalized acreage planted in a particular county. X is a $NT \times K$ matrix of the explanatory variables. β is a $K \times 1$ vector of coefficients reflecting the influence of the explanatory variable matrix on the normalized acres planted, and ε represents the error term, which is assumed to be normally distributed. N represents the total number of counties included in the analysis and T represents the total number of time periods.

The specification in Equation (6) above represents pooled effects; however, fixed effects may be applied. In this case, Equation (6) becomes:

$$\begin{aligned} & y = \rho(I_T \otimes W_N)y + X\beta + X_{INT}\alpha + \varepsilon \\ & \varepsilon \sim N(0, \sigma^2 I_{NT}) \end{aligned} \tag{9}$$

Where X_{INT} is a $N \times N$ county fixed effects matrix.

In the case of a panel data set, we can add in addition to a spatial lag term an AR1 autoregressive term to Equation (9). Equation (9) becomes:

$$y_{it} = \tau y_{it-1} + \rho W_N y_{it} + X_{it}\beta + \alpha_i + \gamma_t + \varepsilon_{it} \tag{10}$$

This is known as a spatial autoregressive (SAR) model for panel data. All spatial models are estimated via MLE routines.

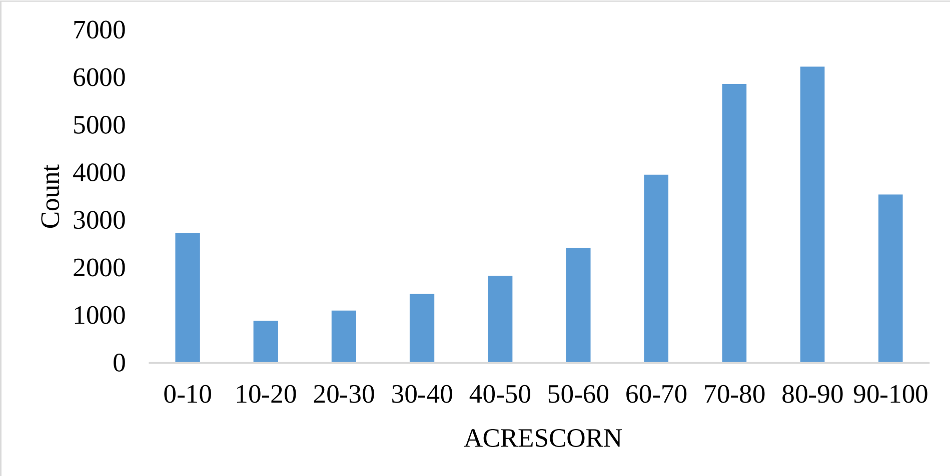
Spatial Expansion Models

Spatial Expansion Models are particularly useful to this analysis because they allow for the linear relationship between an explanatory variable and the dependent variable to differ according to another dimension, such as by space, time, or soil quality. Spatial expansion models can be used to control for spatial heterogeneity in the data by allowing the regression relationship to change as location varies, with local linear fits in clusters of observations that are in close proximity to one another (LeSage and Pace, 2010). Under a spatial expansion specification, first proposed by Cassetti (1972), the effect of an explanatory variable on the variable of interest varies according to latitude and longitude, or by a distance vector. We hypothesize any spatial heterogeneity in our data would be according to latitude, hence a standard Cartesian coordinate expansion is preferred over a distance vector expansion. Furthermore, we perform an expansion by soil quality to test whether the impact of subsidy take and/or expected loss ratios on corn acreage planted varies by soil quality.

Quantile Regression

OLS and spatial methods generate parameter estimates conditional on the mean value of the dependent variable. A quantile regression allows the impact of our covariates on the dependent variable to differ according to the quantile of the dependent variable. An important question for policy makers is whether the impact of insurance availability, subsidization, and pricing on planted acreage differs for county/years that plant a low percentage of their maximum acreage in comparison to county/years that plant close to their maximum acreage. Under the assumptions of OLS, $E[y \mid X] = X\beta$, such that the regression of y on X is the conditional mean, $E[y \mid X]$. Given our dependent variable is not normally distributed (Figure 5), a quantile regression may be more informative.

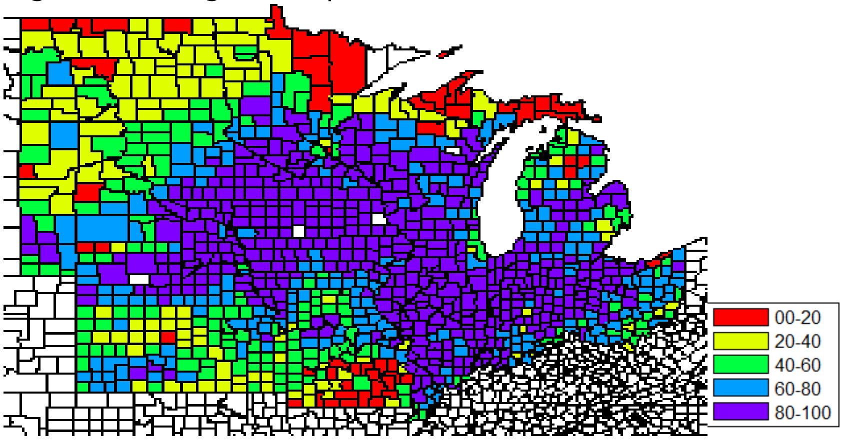
Figure 5. Histogram of ACRESCORN, 1985-2013



Quantile regression, first introduced by Koenker and Bassett (1978), extends this fundamental concept of OLS to the estimation of conditional quantile functions-models in which quantiles of the conditional distribution of the response variable are expressed as functions of our regressors. That is, $\mathrm{E}[\mathrm{y} \mid \mathrm{X}, q] = \mathrm{X}\beta_q$, where q is the quantile to be estimated and $\mathrm{P}(y \leq \mathrm{X}'\beta_q \mid \mathrm{X}) = q, 0 < q < 1$.

To aid in the interpretation of quantile regression results, Figure 6 presents a graph of the average value of *ACRESCORN* for each county over the time series. From the map it is evident that counties in the heart of the Corn Belt consistently plant close to their maximum acreage. Counties that transition toward and away from planting corn occur in the Northwest and Southern areas of the sample space.

Figure 6. Average Value for ACRESCORN, 1985-2013



Variable Filters and Functional Forms

The presence of extreme outliers in a matrix of explanatory variables can lead to substantial distortions and bias in parameter estimates recovered from the model. *CRP*, *ESUBTAKE*, *EYCORN*, *EYSOY*, *ESUBTAKE*, *EGLR*, and *ELR* are winsorized on the basis that each variable has extreme outliers in excel of five standard deviations above their sample means. Each variable is winsorized at the 99.50th percentile. Out of approximately thirty thousand observations, this equates to roughly 60 observations being winsorized. Variables are winsorized rather than trimmed since spatial econometric models require a balanced panel. By winsorizing we avoid dropping counties from the analysis.

As a base case for this empirical analysis we use a standard linear functional form. However, much of the empirical agricultural production literature using double log models as the underlying functional form. A double log model provides for ease of interpretation as the coefficients can be directly interpreted as acreage elasticities. However, a double log model cannot be the true functional form for the model since there are county/years where explanatory variables such as subsidy take and the expected loss ratio are equal to 0. Since a focus of this analysis is to measure the effect of acreage coming into planting with the introduction and subsequent subsidization of crop insurance, changes in energy policy, and changes in expected weather, the true underlying data generating process must be a functional form that allows for variables to be equal to 0 or close to 0. To be somewhat consistent with the previous literature we apply an inverse hyperbolic sine (IHS) variable transformation to both the dependent variable and to the independent variables (excluding the time dummies and fixed effect terms). An IHS transformation is attractive in this case because it is logarithmic but defined at 0. We adopt the notation of Burbridge, Magee, and Robb (1988) and define the IHS transformation to a generic variable x as follows:

$$\sinh^{-1} x = \log(x + \sqrt{x^2 + 1}) \tag{11}$$

Under this variable transformation, the parameter estimates recovered from an econometric model cannot be interpreted as constant elasticity as they would be under a standard logarithmic variable transformation. The derivation of elasticity of a dependent variable Y for a function F with an IHS transformation applied to both the dependent and explanatory variables follows.

Let F be a function where $Y = \alpha + \beta X$. Applying an IHS transformation to F, $\log[Y + (Y^2 + 1)^{1/2}] = \alpha + \beta \log[X + (X^2 + 1)^{1/2}]$. By implicit differentiation, $\frac{dF}{dX} = \frac{dF}{dY}\frac{dY}{dX} = 0$. This implies that $\frac{d(\log[Y + (Y^2 + 1)^{1/2}])}{dX} = \frac{dF}{dY}\frac{dY}{dX} = \frac{d(\log[Y + (Y^2 + 1)^{1/2}])}{dY}\frac{dY}{dX}$. That is, $\frac{dF}{dX} = \frac{1}{(Y^2 + 1)^{1/2}}\frac{dY}{dX}$. Furthermore, $\frac{dF}{dX} = \frac{d(\alpha + \beta \log[X + (X^2 + 1)^{1/2}])}{dX} = \frac{\beta}{(X^2 + 1)^{1/2}}$. Therefore, $\frac{\beta}{(X^2 + 1)^{1/2}} = \frac{1}{(Y^2 + 1)^{1/2}}\frac{dY}{dX}$. Hence, $\frac{dY}{dX} = \beta \frac{(Y^2 + 1)^{1/2}}{(X^2 + 1)^{1/2}}$. By definition, the elasticity of Y with respect to X is $\varepsilon = \frac{dY}{dX}\frac{X}{Y}$. Thus, the elasticity of function F with an IHS variable transformation can be expressed as $\varepsilon = \beta \frac{(Y^2+1)^{1/2}}{(X^2+1)^{1/2}}\frac{X}{Y}$.

Choosing the Functional Form, Specification, and Model

As a base case for our analysis we run OLS and Tobit regression with fixed effects. We hypothesize state or county-level fixed effects are needed as a proxy for aforementioned unobserved variables in our data. A Hausman Test will allow us to test our hypothesis of fixed effects in lieu of a pooled effects model. There are several qualitative and statistical methods that we will use to aid us in determining (1) the best functional form (linear, IHS), (2) the best specification (OLS, Tobit, Spatial, Expansion, or Quantile), and (3) the best model (M1-M7).

Within sample, there are several quantitative and qualitative ways to select the superior functional form, specification, and model. If the spatial lag term recovered from the spatial models is positive and statistically significant, this indicates a spatial specification is preferred to a Tobit or OLS specification in order to control for spatial dependency and heterogeneity present in the dataset. Furthermore, a Moran's I Test can be used to test for spatial autocorrelation of the residuals from an OLS regression. If the Moran's I-Statistic is greater than 1.96 this indicates rejection of the null hypothesis (H0: No spatial autocorrelation of OLS regression residuals). In this case a spatial specification would be appropriate in order to manage issues of spatial dependency and/or heterogeneity in the data. The Akaike Information Criterion (AIC) can be used to determine the best specification, and a Likelihood Ratio Test can be used to choose between nested models once we have selected the best functional form and specification.

Cross-validation will be the out of sample statistical method used to decide between functional forms, specifications, and models. For each functional form, specification, and model of interest, we cross-validate the model by iterating from i = 1985 to 2013, and estimate the model for all but the ith year. Parameter estimates generated over each iteration are used to calculate fitted values of the dependent variable in the ith year. The model with the lowest sum of squared errors is preferred.

Empirical Models

The generic model for this analysis can be expressed as:

$$y_{it} = \tau y_{it-1} + \rho W_N y_{it} + X_{it}\beta + X_{it-1}\gamma + T_t\delta + \alpha_i + \varepsilon_{it} \tag{12}$$

where y in county i in time t is a function of an autoregressive term, τy_{it-1}, a spatial lag term, $\rho W_N y_{it}$, a matrix of explanatory variables in time t, X_{it}, a matrix of explanatory variables observed in time t-1, X_{it-1}, a matrix of time dummies, T_t, a time invariant individual fixed effect for each county i, α_i, and an error term, ε_{it}. A summary of the seven different empirical models estimated is presented in Table 5.

Model M1 is the base model for this analysis. It includes lagged acreage, expected yield performance and yield risk for soybeans and corn, the relative price of soybeans to corn, a price to cost ratio, urbanization, and ethanol demand. These variables have been identified in the previous literature as primary acreage determinants. CRP enrollment as a percentage of county acres is added to the base model in M2, along with a matrix of time dummies to reflect structural changes to the FCIP which could result in changes to planted acreage, such as the introduction of revenue and group insurance, changes to subsidies, or changes in program target loss ratios. Expected yield performance for corn and soybeans are dropped in M3 in order to compare the robustness of results between M2 and M3. In M4 the subsidy-adjusted expected loss ratio is added in order to distinguish between the effect insurance availability has on planted acreage and acreage response specifically due to subsidies and expected returns on insurance. Expected weather is added to the covariates in M5. In M6 the expected subsidy-adjusted loss ratio is replaced by expected gross loss ratio and expected subsidy take. In M7 expected weather terms are added to the covariates in M6.

Table 5. Summary of Empirical Models

M1	*LAGCORN, LAGSOY, EYIELDCORN, EYIELDSOY, CVCORN, CVSOY, PXRATIO, PCRATIO, URBAN, ETHANOL*
M2	*LAGCORN, LAGSOY, CRP, EYIELDCORN, EYIELDSOY, CVCORN, CVSOY, PXRATIO, PCRATIO, URBAN, ETHANOL, TIME*
M3	*LAGCORN, LAGSOY, CRP, CVCORN, CVSOY, PXRATIO, PCRATIO, ETHANOL, TIME*
M4	*LAGCORN, LAGSOY, CVCORN, CVSOY, PXRATIO, PCRATIO, ETHANOL, ELR, TIME*
M5	*LAGCORN, LAGSOY, CVCORN, CVSOY, PXRATIO, PCRATIO, ETHANOL, ELR, ETEMP,* $ETEMP^2$*, EPREC, ETEMP*EPREC, TIME*
M6	*LAGCORN, LAGSOY, CVCORN, CVSOY, PXRATIO, PCRATIO, ETHANOL, EGLR, ESUBTAKE, TIME*
M7	*LAGCORN, LAGSOY, CVCORN, CVSOY, PXRATIO, PCRATIO, ETHANOL, EGLR, ESUBTAKE, ETEMP,* $ETEMP^2$*, EPREC, ETEMP*EPREC, TIME*

The models are estimated using a spatial autoregressive model for panel data with county-level fixed effects, a spatial (latitudinal) expansion model with county-level fixed effects, and a quantile regression with state-level fixed effects. Type 1 Tobit and OLS models with state-level fixed effects are used as base specifications for comparison.

CHAPTER V

RESULTS

OLS, Tobit, and Spatial Lag Results

A Hausman Test statistic indicates rejection of the null hypothesis of consistent parameter estimates using a pooled panel model, supporting our theoretical justification that fixed effects are required to control for unobservable heterogeneity present in the data.

Table 6. Regression Results, OLS with State-Level Fixed Effects, M1-M7

Variable	*M1*	*M2*	*M3*	*M4*	*M5*	*M6*	*M7*
LAGCORN	0.832***	0.819***	0.836***	0.835***	0.834***	0.833***	0.833***
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
LAGSOY	0.036***	0.039***	0.047***	0.047***	0.056***	0.047***	0.056***
	(0.003)	(0.002)	(0.002)	(0.002)	(0.003)	(0.002)	(0.003)
CRP		0.025*	0.045***				
		(0.015)	(0.015)				
EYCORN	0.029***	0.035***					
	(0.004)	(0.004)					
EYSOY	-0.032**	0.003					
	(0.014)	(0.014)					
CVCORN	-1.605***	-2.273***	-5.288***	-5.287***	-4.893***	-5.639***	-5.263***
	(0.570)	(0.566)	(0.495)	(0.526)	(0.531)	(0.540)	(0.544)
CVSOY	-1.487**	-0.188	-1.247*	-1.294**	-0.280	-1.357*	-0.385
	(0.655)	(0.656)	(0.639)	(0.642)	(0.644)	(0.653)	(0.655)
PXRATIO	-5.198***	-9.233***	-9.015***	-9.040***	-9.080***	-9.206***	-9.312***
	(0.267)	(0.385)	(0.386)	(0.386)	(0.389)	(0.387)	(0.391)
PCRATIO	2.673***	0.614*	0.997***	1.051***	1.081***	0.996***	1.082***
	(0.211)	(0.357)	(0.355)	(0.355)	(0.357)	(0.356)	(0.358)
URBAN	0.286***	0.888***					
	(0.051)	(0.062)					
ETHANOL	0.286***	0.008	0.027***	0.028***	0.022***	0.034***	0.027***
	(0.002)	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)
ELR				0.005	-0.039		
				(0.045)	(0.045)		
ESUBTAKE						12.044***	12.096***
						(1.125)	(1.124)
EGLR						0.094	0.019
						(0.091)	(0.091)
ETEMP					0.530		0.297
					(0.598)		(0.605)
ETEMP^2					0.007		0.014
					(0.012)		(0.012)
EPREC					0.074***		0.076***
					(0.015)		(0.015)
*ETEMP*EPREC*					-0.004***		-0.004
					(0.000)		(0.000)
n	26088	26088	26088	26045	26045	25462	25462
σ^2	72.953	70.427	71.296	71.138	70.491	70.049	69.413
adj. r^2	0.846	0.852	0.849	0.850	0.851	0.847	0.845

*Note: Fixed Effects and Time Dummies are omitted from reported results. Standard Errors in parentheses. *, **, and *** indicates significance at α = 0.01, 0.05, and 0.10 respectively.*

Table 7. Regression Results, Tobit with State-Level Fixed Effects, M1-M7

Variable	*M1*	*M2*	*M3*	*M4*	*M5*	*M6*	*M7*
LAGCORN	0.839***	0.825***	0.842***	0.841***	0.839***	0.840***	0.838***
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
LAGSOY	0.035***	0.039***	0.047***	0.047***	0.055***	0.047***	0.055***
	(0.003)	(0.003)	(0.002)	(0.003)	(0.003)	(0.003)	(0.003)
CRP		0.025	0.046***				
		(0.015)	(0.015)				
EYCORN	0.029***	0.035					
	(0.004)	(0.004)					
EYSOY	-0.040***	-0.004					
	(0.014)	(0.014)					
CVCORN	-1.685***	-2.415***	-5.233***	-5.244***	-4.892***	-5.622***	-5.262***
	(0.582)	(0.578)	(0.509)	(0.537)	(0.537)	(0.551)	(0.550)
CVSOY	-1.518**	-0.179	-1.171*	-1.191*	-0.277	-1.228**	-0.383
	(0.669)	(0.669)	(0.652)	(0.655)	(0.657)	(0.666)	(0.667)
PXRATIO	-5.323***	-9.625***	-9.355***	-9.377***	-9.085***	-9.404***	-9.317***
	(0.278)	(0.393)	(0.394)	(0.394)	(0.396)	(0.396)	(0.398)
PCRATIO	2.701***	0.337	0.767**	0.820**	1.083***	0.831***	1.083***
	(0.216)	(0.361)	(0.362)	(0.362)	(0.364)	(0.364)	(0.365)
URBAN	0.289***	0.909***					
	(0.058)	(0.064)					
ETHANOL	0.011***	0.020***	0.039***	0.040***	0.025***	0.046***	0.030***
	(0.002)	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)
ELR				0.005	-0.045		
				(0.046)	(0.046)		
ESUBTAKE						12.724***	12.098***
						(1.154)	(1.149)
EGLR						0.097	0.020
						(0.093)	(0.092)
ETEMP					0.526		0.293***
					(0.557)		(0.564)
ETEMP^2					0.007		0.014
					(0.012)		(0.012)
EPREC					0.074***		0.075***
					(0.000)		(0.000)
					-0.004***		-0.004***
*ETEMP*EPREC*					(0.000)		(0.000)
n	26088	26088	26088	26045	26045	25462	25462
σ^2	75.720	73.012	73.939	73.774	73.084	72.692	72.014
loglik	-68528	-68044	-68207	-68065	-67941	-66340	-66218

*Note: Fixed Effects and Time Dummies are omitted from reported results. Standard Errors in parentheses. *, **, and *** indicates significance at α = 0.01, 0.05, and 0.10 respectively.*

Tobit and OLS results are presented in Tables 6 and 7. OLS results do not appear to be significantly biased compared to Tobit results. A Moran I Test (Table 8) supports the use of a spatial lag specification over an OLS specification. With Moran I-Test statistics exceeding 1.96 for each model (M1-M7) in each year we reject the null hypothesis of the Moran I Test (H0:

There is no spatial correlation in the residuals of an OLS regression model) and confirm our theoretical justification that planted acreage models should account for spatial relationships.

The results for models M1-M7 estimated using a spatial lag specification with county-level fixed effects is presented in Table 9. Both *LAGCORN* and *LAGSOY* are statistically significant and positive in all models. While the magnitude of the parameter estimate for *LAGSOY* is approximately the same under both a spatial and OLS specification, the parameter estimate for *LAGCORN* generated under a spatial specification is about half the magnitude as under an OLS. *CRP* is surprisingly positive and not statistically significant, possibly because it represents a very small portion of overall acreage in the Corn Belt. As hypothesized, *EYCORN* is positive and statistically significant. *EYSOY* is negative and statistically significant, meaning for decreases in expected yields for soybeans farmers may abandon crop rotation practices and plant more corn. *CVCORN* and *CVSOY* are statistically insignificant in models that explicitly include expected yields (M1-M2). However, when *EYCORN* and *EYSOY* are dropped from the model, the signs of *CVCORN* and CVSOY are robust, the parameters are statistically significant, and the effect of *EYCORN* and *EYSOY* is absorbed by the *CVCORN* and *CVSOY* terms.

PXRATIO is negative and statistically significant. That is, for an increase in soybean prices relative to corn, farmers abandon crop rotation in the short-term and plant less corn. *PCRATIO* is positive and statistically significant, indicating for increases in prices received relative to the price of inputs, farmers will plant increased corn.

URBAN is negative as expected but statistically insignificant. Urbanization may be statistically insignificant because data is only available at the state-level. *ETHANOL* is surprisingly *negative* and significant. That is, for mean *ACRESCORN*, an increase in lagged ethanol consumption results in a decrease in indexed corn planted acreage, which is counterintuitive. We test the robustness of this result across different quantiles of the dependent variable later in this analysis.

Table 8. Moran I-Test Statistics for M1-M7

Year	*M1*	*M2*	*M3*	*M4*	*M5*	*M6*	*M7*
1985	6.635	6.635	6.635	7.175	6.488	6.488	6.488
1986	12.876	11.935	12.408	13.646	10.619	10.619	10.619
1987	11.427	11.388	12.589	12.549	11.301	11.301	11.301
1988	13.239	13.183	17.259	18.560	12.113	12.113	12.113
1989	16.388	16.165	16.498	17.268	17.224	16.129	16.055
1990	19.596	19.482	21.071	20.842	15.829	19.615	14.161
1991	13.631	13.354	20.845	20.255	10.494	18.607	9.857
1992	14.670	14.385	15.956	15.711	11.981	14.652	10.709
1993	11.150	10.321	13.040	13.362	11.133	11.506	10.279
1994	11.791	11.926	13.836	13.927	10.284	13.812	10.539
1995	26.039	24.418	25.147	24.607	21.715	22.319	20.057
1996	21.074	20.837	21.197	21.322	20.048	20.006	17.545
1997	18.905	18.848	22.279	21.250	17.187	20.338	15.464
1998	6.672	6.714	13.209	9.874	9.700	10.857	10.706
1999	9.226	9.151	10.241	9.902	7.186	9.826	7.447
2000	15.012	13.824	16.931	16.956	15.537	15.430	13.188
2001	15.654	13.488	13.942	16.462	13.689	13.578	11.102
2002	19.658	17.354	17.712	18.922	16.239	18.563	15.113
2003	13.400	13.390	13.648	13.746	11.501	12.108	10.308
2004	16.461	15.712	20.043	20.733	16.361	18.861	14.571
2005	14.176	13.859	15.983	15.169	14.622	13.478	13.298
2006	19.043	19.077	19.970	17.747	14.891	16.988	14.701
2007	13.790	13.854	16.373	16.468	13.154	13.958	11.263
2008	17.163	17.372	20.503	20.458	15.972	17.719	13.026
2009	8.923	7.941	8.230	9.519	10.065	6.530	7.118
2010	15.525	15.538	17.473	17.565	10.782	15.945	10.238
2011	18.225	17.704	20.637	21.174	16.443	16.770	13.981
2012	17.800	18.016	18.611	19.173	18.621	16.640	16.336
2013	23.183	23.262	23.104	23.096	20.375	21.471	18.906

ELR is positive and statistically significant. An increase in expected loss ratios net of subsidies results in a marginal increase in *ACRESCORN.* This is in contrast to the parameter estimates recovered under Tobit and OLS which were *negative*. Similarly, *ESUBTAKE* and

EGLR are positive and statistically significant indicating increases in expected subsidy take and expected gross loss ratios result in increased corn planting.

Table 9. Regression Results, Spatial Lag with County-Level Fixed Effects, M1-M7

Variable	*M1*	*M2*	*M3*	*M4*	*M5*	*M6*	*M7*
LAGCORN	0.402***	0.401***	0.403***	0.407***	0.412***	0.381***	0.389***
	(0.005)	(0.005)	(0.005)	(0.004)	(0.004)	(0.005)	(0.005)
LAGSOY	0.025***	0.027***	0.021***	0.021***	0.020***	0.018***	0.017***
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
CRP		0.011	0.009				
		(0.029)	(0.028)				
EYCORN	0.035***	0.044***					
	(0.006)	(0.006)					
EYSOY	-0.182***	-0.152***					
	(0.021)	(0.022)					
CVCORN	-0.980	-1.012	-2.409***	-2.919***	-2.893*	-2.346***	-2.286***
	(0.636)	(0.646)	(0.594)	(0.629)	(0.627)	(0.682)	(0.680)
CVSOY	0.184	0.788	1.954*	1.927**	2.112***	1.612*	1.769**
	(0.759)	(0.787)	(0.771)	(0.770)	(0.769)	(0.830)	(0.831)
PXRATIO	-1.973***	-3.028***	-2.930***	-3.011***	-3.187***	-2.789***	-3.061***
	(0.205)	(0.293)	(0.292)	(0.289)	(0.294)	(0.300)	(0.310)
PCRATIO	1.149***	1.364***	1.315***	1.373***	1.555***	1.171***	1.346***
	(0.116)	(0.263)	(0.262)	(0.259)	(0.262)	(0.267)	(0.270)
URBAN	-0.064	-0.044					
	(0.047)	(0.051)					
ETHANOL	-0.003**	-0.021***	-0.022***	-0.022***	-0.035***	-0.018***	-0.030***
	(0.001)	(0.005)	(0.004)	(0.004)	(0.005)	(0.004)	(0.005)
ELR				0.096**	0.098**		
				(0.039)	(0.039)		
ESUBTAKE						4.552***	4.220***
						(0.928)	(0.932)
EGLR						0.181**	0.185**
						(0.078)	(0.078)
ETEMP					3.255***		2.926***
					(0.737)		(0.792)
$ETEMP^2$					-0.052***		-0.041***
					(0.014)		(0.015)
EPREC					0.071***		0.080***
					(0.014)		(0.019)
*ETEMP*EPREC*					-0.004***		-0.005***
					(0.000)		(0.000)
n	19024	19024	19024	19024	19024	16385	16385
σ^2	27.409	27.287	27.394	27.433	27.221	25.055	24.922
loglik	-59379	-59338	-59369	-59377	-59295	-50479	-50401
adj. r^2	0.908	0.909	0.908	0.908	0.909	0.899	0.900
ρ	0.655***	0.653***	0.653***	0.647***	0.643***	0.675***	0.664***

*Note: Fixed Effects and Time Dummies are omitted from reported results. Standard Errors in parentheses. *, **, and *** indicates significance at α = 0.01, 0.05, and 0.10 respectively.*

Finally, *ETEMP* and *EPREC* are statistically significant factors driving planted acreage. As hypothesized there is a nonlinear relationship between *ETEMP* and *ACRESCORN*, and *ETEMP* and *EPREC* are jointly considered by the farmer when allocating acreage to corn. Overall, likelihood ratio tests for nested models suggest the more complex models M5 and M7 are preferred to the more parsimonious models M1-M4 and M6.

Spatial Expansion Model Results

The results of models M5 and M7 with a spatial lag specification and a latitudinal expansion term applied to *ETEMP* and $ETEMP^2$ are presented in Table 10. Parameter estimates generated under a spatial lag specification remain robust with the addition of the expansion terms with the exception of *ETEMP* and *EPREC.*

Holding all other variables in the regression constant, and evaluating at mean precipitation, the effect of an increase in *ETEMP* from 22 to 23 degrees Celsius on acreage is presented for three different latitudes in Table 11. In general, counties further north in the sample appear to be more responsive to an increase in expected temperature. Under M5, an increase in *ETEMP i*n the southernmost region of the sample results in a decrease to planted acreage. This has interesting implications under future climate change scenarios. In the more northern region of the sampling an increase in temperature from 22 to 23 degrees Celsius results in a 2% increase to *ACRESCORN*. The results from the latitudinal expansion of M7 do not seem to make any sense or result in feasible changes to *ACRESCORN*.

Models M5 and M7 were estimated with a soil expansion applied to *ELR*, *ESUBTAKE*, and *EGLR* to test whether farmer acreage response to subsidy take and/or expected loss ratios varies according to soil quality. Results are reported in Table 12. The expansion of *ELR*, *ESUBTAKE*, and *EGLR* by *SOIL* is not statistically significant. Furthermore, after applying a soil expansion to these terms the parameter estimates for *ELR*, *ESUBTAKE*, and *EGLR* are no longer statistically significant.

Table 10. Regression Results, Spatial Lag with Latitudinal Expansion and County-Level Fixed Effects, M5 and M7

Variable	*M5*	*M7*
LAGCORN	0.416***	0.391***
	(0.004)	(0.005)
LAGSOY	0.021***	0.017***
	(0.003)	(0.003)
CVCORN	-2.695***	-2.135
	(0.626)	(0.679)
CVSOY	2.339***	2.098**
	(0.771)	(0.832)
PXRATIO	-3.497***	-3.280***
	(0.295)	(0.311)
PCRATIO	1.438***	1.244***
	(0.262)	(0.270)
ETHANOL	-0.029***	-0.025***
	(0.005)	(0.005)
ELR	0.093**	
	(0.039)	
ESUBTAKE		1.244***
		(0.930)
EGLR		-0.025***
		(0.078)
ETEMP	-21.135***	4.374***
	(0.896)	(0.542)
*ETEMP*LAT*	0.421***	0.173**
	(0.021)	(0.018)
ETEMP^2	0.276***	-19.993**
	(0.023)	(0.022)
ETEMP^2**LAT*	-0.004***	0.395***
	(0.000)	(0.000)
EPREC	0.034*	0.263***
	(0.018)	(0.019)
*ETEMP*EPREC*	-0.002***	-0.004***
	(0.000)	(0.000)
n	19024	16385
σ^2	27.213	24.848
loglik	-59252	-50365
adj. r^2	0.909	0.900
ρ	0.634***	0.663***

*Note: Fixed Effects and Time Dummies are omitted from reported results. Standard Errors in parentheses. *, **, and *** indicates significance at α = 0.01, 0.05, and 0.10 respectively.*

Table 11. Change in ACRESCORN Given Changes in Temperature and Latitude

	35 North		*40 North*		*45 North*	
	M5	*M7*	*M5*	*M7*	*M5*	*M7*
Δ *ACRESCORN, T=22 to 23*	-0.280	-267.131	0.925	-177.391	2.130	-87.651

The parameter estimate for *ELR*SOIL* is positive, meaning counties with higher soil quality plant more corn with increases in the expected loss ratio net of subsidies. When the effect of subsidies and expected loss ratios is decomposed in M7, the results suggest that counties with high soil quality are most responsive to increases in expected subsidy take, whereas counties with high soil quality decrease planting when there is an increase in the expected *gross* loss ratio.

Overall, a soil expansion around *ELR*, *ESUBTAKE*, and *EGLR* does not appear to be the correct specification. It could be that the measure of soil quality used is confounding. The NCCPI soil quality indicator considers county yield performance and climate change, which is already directly captured in the model. In future research and development of the model we will try expanding across different measures of soil quality, such as organic carbon content or water content.

Table 12. Regression Results, Spatial Lag with Soil Expansion and County-Level Fixed Effects, M5 and M7

Variable	*M5*	*M7*
LAGCORN	0.411***	0.389***
	(0.004)	(0.0056)
LAGSOY	0.020***	0.017***
	(0.003)	(0.003)
CVCORN	-2.895***	-2.266***
	(0.627)	(0.681)
CVSOY	2.112***	1.755**
	(0.770)	(0.831)
PXRATIO	-3.171***	-3.043***
	(0.294)	(0.310)
PCRATIO	1.556***	1.349***
	(0.262)	(0.270)
ETHANOL	-0.035***	-0.029***
	(0.005)	(0.005)
ELR	0.011	
	(0.122)	
*ELR*SOIL*	0.002	
	(0.002)	
ESUBTAKE		4.571
		(4.169)
*ESUBTAKE*SOIL*		0.390
		(0.242)
EGLR		-0.005
		(0.063)
*EGLR*SOIL*		-0.004
		(0.004)
ETEMP	3.247***	2.939***
	(0.737)	(0.792)
ETEMP^2	-0.052***	-0.042***
	(0.014)	(0.015)
EPREC	0.071***	0.080***
	(0.018)	(0.019)
*ETEMP*EPREC*	-0.004***	-0.005***
	(0.000)	(0.000)
n	19024	16385
σ^2	27.202	24.911
loglik	-59280	-50406
adj. r^2	0.909	0.899
ρ	0.645***	0.665***

*Note: Fixed Effects and Time Dummies are omitted from reported results. Standard Errors in parentheses. *, **, and *** indicates significance at α = 0.01, 0.05, and 0.10 respectively.*

Regression Results Using Transformed Variables

An IHS variable transformation is applied to the dependent and explanatory variables (with the exception of time dummies and fixed effect terms) and M5 and M7 are re-estimated under a spatial lag specification, and M5 under a spatial lag specification with a latitudinal expansion applied to *ETEMP* and *ETEMP*^{2}. The results of these regressions are reported in Table 13.

Table 13. Regression Results, Spatial Lag with IHS Variable Transformation and County-Level Fixed Effects, M5, M5 Expansion, and M7

Variable	*M5*	*M5 Expansion*	*M7*
IHS(LAGCORN)	0.503***	0.504***	0.451***
	(0.004)	(0.004)	(0.005)
IHS(LAGSOY)	0.008***	0.008***	0.007***
	(0.002)	(0.002)	(0.002)
IHS(CVCORN)	-0.072***	-0.071***	-0.061
	(0.012)	(0.012)	(0.012)
IHS(CVSOY)	0.041***	0.047***	0.045**
	(0.014)	(0.014)	(0.015)
IHS(PXRATIO)	-0.159***	-0.162***	-0.135***
	(0.014)	(0.014)	(0.014)
IHS(PCRATIO)	0.012*	0.011	0.006
	(0.007)	(0.007)	(0.007)
IHS(ETHANOL)	-0.003	-0.000	0.000
	(0.004)	(0.004)	(0.004)
IHS(ELR)	0.007***	0.007***	
	(0.002)	(0.002)	
IHS(ESUBTAKE)			0.087***
			(0.016)
IHS(EGLR)			0.008***
			(0.002)
IHS(ETEMP)	5.433***	28.867	5.237***
	(0.337)	(20.743)	(0.246)
*IHS(ETEMP)*IHS(LAT)*		-7.170	
		(4.479)	
*IHS(ETEMP)*IHS(ETEMP)*	-0.455***	-5.556**	-0.369***
	(0.072)	(2.564)	(0.070)
IHS(ETEMP) IHS(ETEMP)*IHS(LAT)*		1.377**	
		(0.550)	
IHS(EPREC)	1.168***	0.693**	1.437***
	(0.085)	(0.347)	(0.087)
*IHS(ETEMP)*IHS(EPREC)*	-0.336***	-0.209**	-0.401***
	(0.023)	(0.091)	(0.022)
n	19024	19024	16385
σ^2	0.009	0.009	0.007
loglik	17087	17110	16284
adj. r^2	0.926	0.926	0.916
ρ	0.557***	0.556***	0.621***

*Note: Fixed Effects and Time Dummies are omitted from reported results. Standard Errors in parentheses. *, **, and *** indicates significance at α = 0.01, 0.05, and 0.10 respectively.*

Parameter signs are mostly robust between the linear and IHS results with a few notable exceptions. *ETEMP* and *ETEMP*LAT* are no longer statistically significant under M5 with an expansion. *ETHANOL* is no longer statistically significant, and in the case of M7 becomes positive. Spatial ρ's are slightly smaller after applying the IHS variable transformation.

Unlike the double log functional form which results in a constant elasticity between X and Y, under both linear and IHS functional forms the elasticity of *ACRESCORN* to a particular covariate will vary according to both the value of the covariate as well as the value of *ACRESCORN*. Furthermore, unlike a linear model, under an IHS variable transformation applied to both the dependent and explanatory variables, there is no longer a linear marginal effect of each covariate on the dependent variable. Graphs will help to illustrate this result. Point elasticities for key explanatory variables of interest are graphed in Figures 7 through 10. Since observed X's and fitted values for *ACRESCORN* for the corresponding observation are used to calculate the elasticity, for some values of X we recover multiple elasticities, as the fitted values for *ACRESCORN* depend on the observed value of the other covariates in the model. In each figure the elasticities calculated under the linear function form is presented on the left, and the elasticites calculated under the IHS functional form follows to the right. The interquartile range (IQR) of observed values of the covariate is denoted by vertical red lines.

Acreage response to changes in *ELR* is positive and inelastic under both linear and IHS functional forms. Furthermore, in both cases, changes in *ELR* result in a greater acreage response for higher observed values of *ELR*, although levels are overall marginal. Similarly, acreage response to *SUBTAKE* and *EGLR* are inelastic but with the largest response occurring for changes at the highest observed values of the covariates. The largest difference in estimated acreage elasticites between the two functional forms occurs for *ETEMP*. Under a linear functional form, point elasticities for *ACRESCORN* are large and negative for changes in *ETEMP* for the highest values of *ETEMP*. For the lowest levels, values are positive, in line with our hypothesis. Under an IHS variable transformation, however, acreage response to changes in *ETEMP* is more inelastic than under a linear functional form.

Figure 7. Comparison of ELR vs. Elasticity of ACRESCORN, M5 Linear vs. M5 IHS

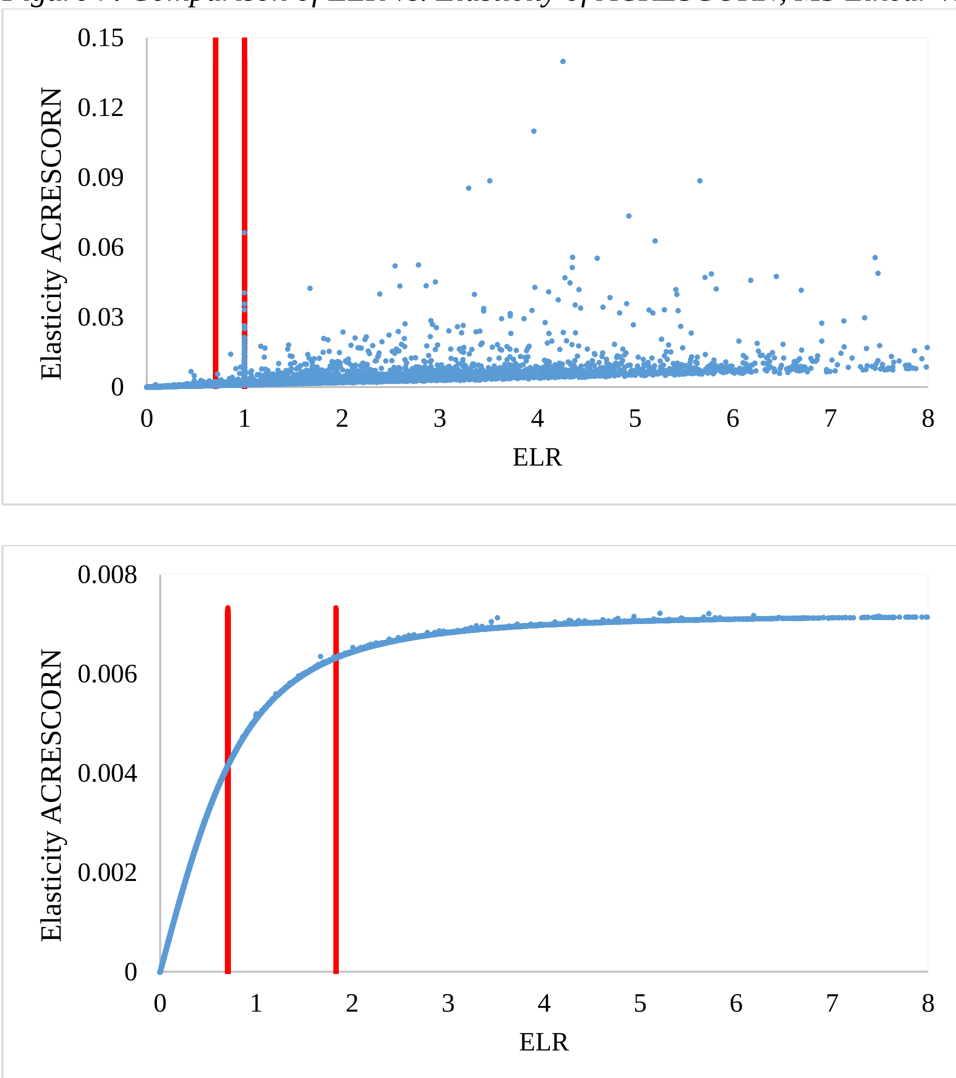
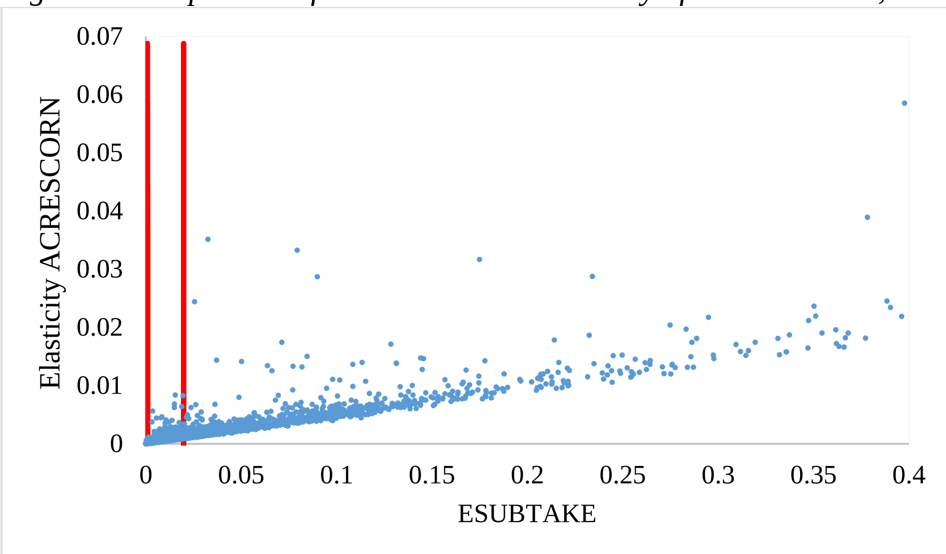


Figure 8. Comparison of ESUBTAKE vs. Elasticity of ACRESCORN, M7 Linear vs. M7 IHS



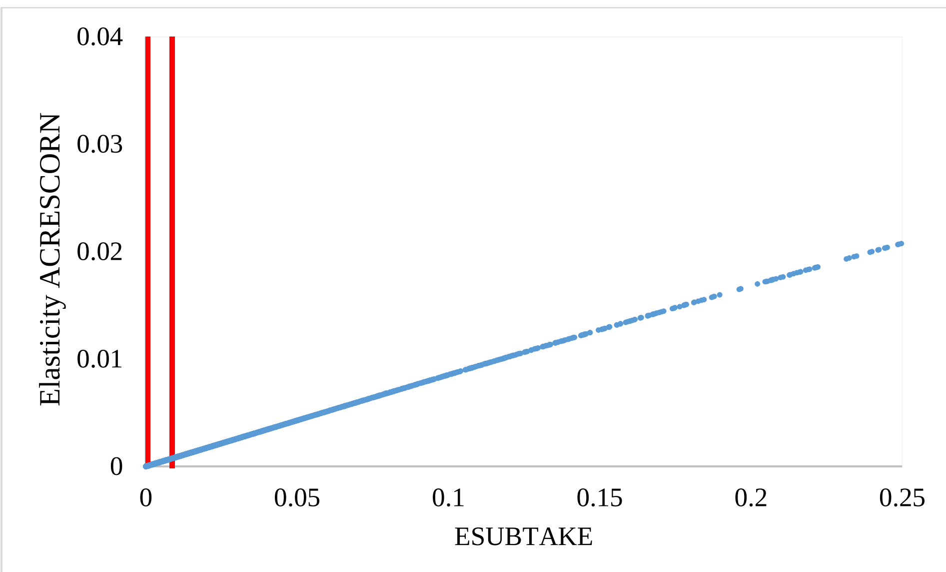
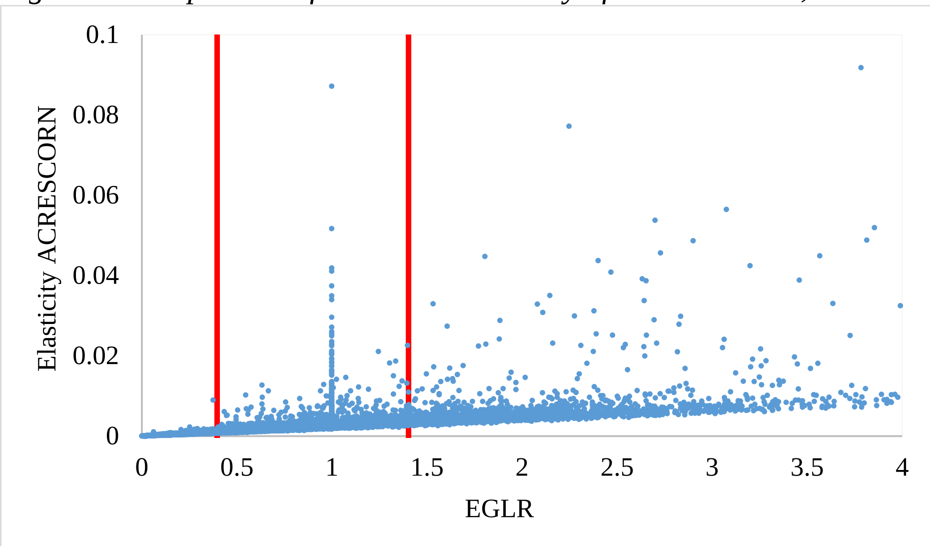


Figure 9. Comparison of EGLR vs. Elasticity of ACRESCORN, M7 Linear vs. M7 IHS



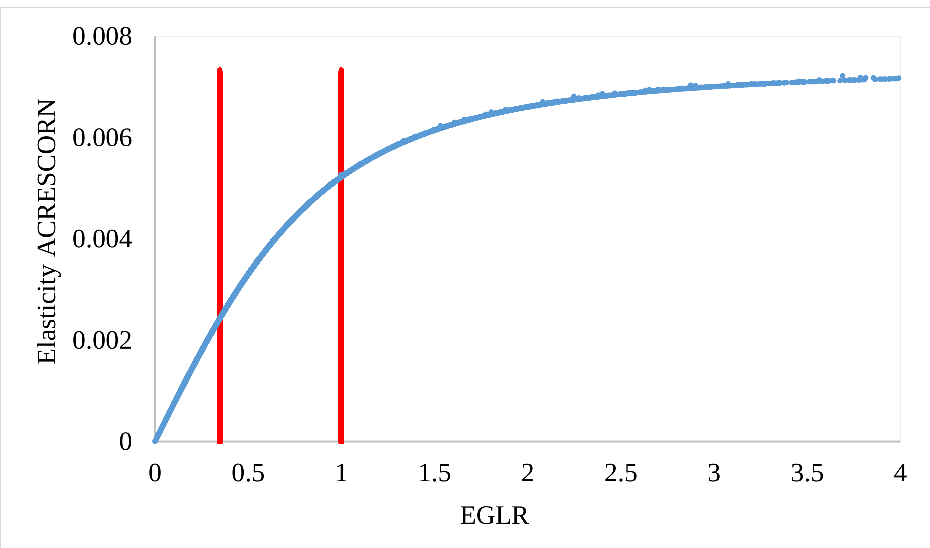
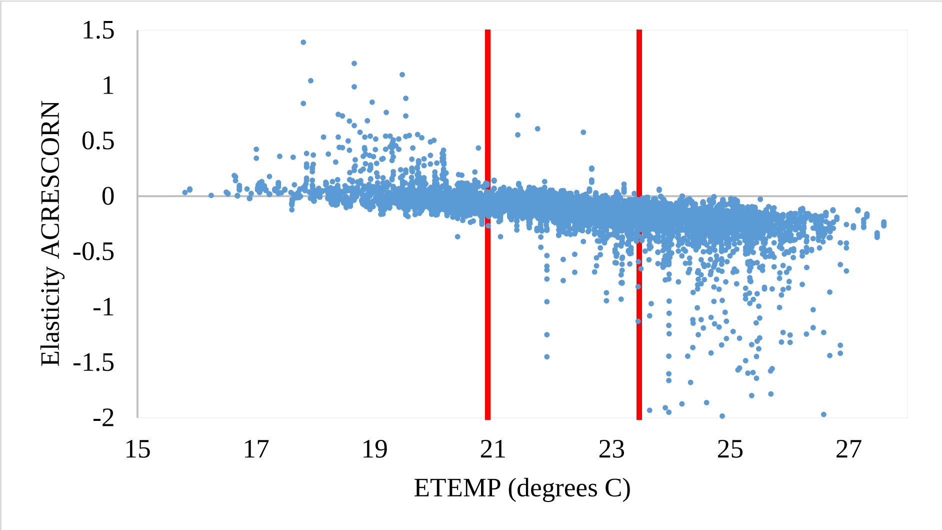
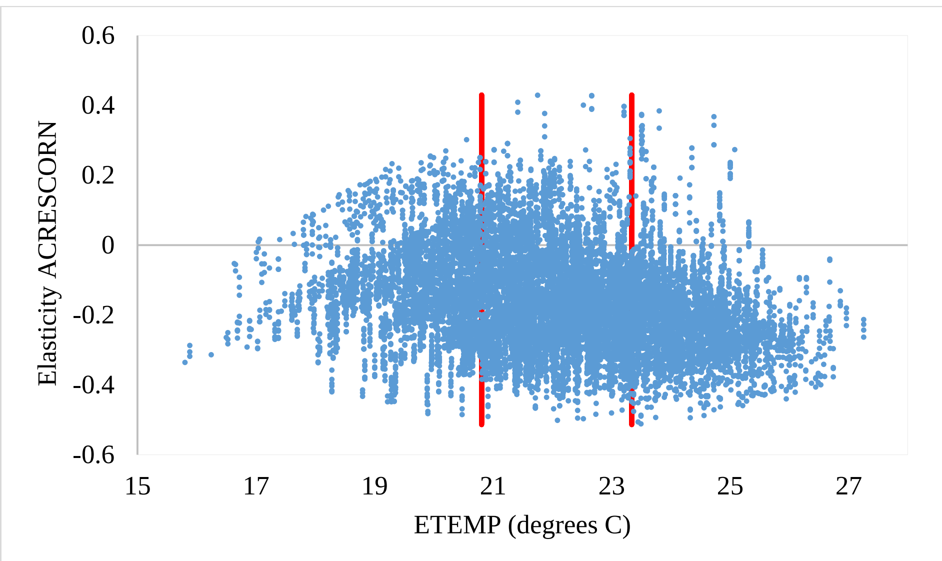


Figure 10. Comparison of ETEMP vs. Elasticity of ACRESCORN, M5 Linear vs. M5 IHS





Model Validation and Selection

Likelihood ratio tests for nested models support the use of M5 and M7 compared to more parsimonious models. Candidate specifications include spatial lag and spatial lag with a latitudinal expansion term applied to *ETEMP* and $ETEMP^2$. Candidate functional forms are linear and IHS. In-sample test statistics such as the AIC can be used to compare the various models, specifications, and functional forms. Furthermore, out of sample methods such as cross-validation can be used. The AICs and sum of squared errors from cross-validation for the various candidate models, specifications, and functional forms is presented in Table 14.

Table 14. In and Out of Sample Model Validation

	AIC	*Cross-Validation SSE*	*n*
M5	119,935	3,859,149,679	19024
M5-IHS	-32,830	5,270,233	19024
M5 Expansion	117,143	3,847,520,487	19024
M5 Expansion-IHS	-35,574	5,246,870	19024
M7	101,979	13,393,001,782	16385
M7-IHS	-31,388	6,019,062	16385

The combination of models, functional forms, and specifications with the lowest AIC in absolute terms is preferred. This would mean M7 with an IHS variable transformation would be a good candidate for the true underlying relationship between the covariates and indexed acreage. M5 with an IHS variable transformation, and M5 with a temperature expansion and IHS variable transformation also appear to be good candidates. Since we would like to use the model to predict changes in planted acreage given structural changes in the FCIP, an out of sample test such as cross-validation is used to see how each specification performs out of sample. Performance out of sample varies widely, with the M5 with IHS, M5 Expansion with IHS, and M7 with IHS performing the strongest. Due to the strong performance both in and out of sample, and because of the demonstrated relationship between expected temperature and indexed acreage, we select M5 with an IHS variable transformation and latitudinal expansion term as the complete model for the analysis.

Quantile Regression Results

Under OLS, Tobit, and Spatial econometric specifications we are able to estimate the impact of our covariates conditional on the *mean* value of *ACRESCORN*, but we cannot determine whether the impact of our key covariates of interest varies in county/years where corn planted acreage is close to its maximum over the time series (e.g. for county/years where *ACRESCORN* is near

100), versus county/years which are transitioning toward or away from planting corn over the time series (e.g. for county/years where *ACRESCORN* is near 0). A quantile regression is used to analyze the robustness of parameter estimates across various quantiles of the dependent variable.

To assist us in interpreting these quantile regressions, it is necessary to understand the counties that are likely to fall in the lower quantiles and those that fall in the higher quantiles. Recall Figure 3, a map of the average *ACRESCORN* observed in each county over the complete time series for the sample counties. Notably, the regions to the north and northwest of the sample space (namely northern Minnesota, North and South Dakota) will likely fall in the lower quantiles, as well as a smaller pocket of counties in the southern portion of the sample.

Recall the histogram of *ACRESCORN* in Figure 5 showing the distribution of observed values is not normal but left-skewed. Since the distribution of our dependent variable is not normal, regression results generated at the mean (OLS) might not be particularly insightful, and the results of a quantile regression, in particular the interpretation of explanatory variables with parameter estimates that vary widely according to quantile, can surely serve to provide more clarity as to the true underlying relationship between the covariates and *ACRESCORN*. A key limitation of the quantile regression is that it does not account for spatial dependency present in the data. So the quantile regression is used in this analysis more as an interpretive tool for covariates with parameter estimates that are not robust across the different quantiles of the dependent variable. The results of a quantile regression for M5 with a latitudinal expansion of *ETEMP* and an IHS variable transformation with state-level fixed effects is presented in Table 15.

LAGCORN, *LAGSOY*, *CVCORN*, *CVSOY*, and *PCRATIO* are mostly robust across quantiles. *PXRATIO* increases but remains negative for higher quantiles, indicating relative prices have a greater effect on counties when they are not planting near their maximum observed acreage. This result suggests that relative prices are a driver of increased acreage. *ETHANOL* increases by quantile, meaning increases in ethanol demand have caused counties planting a high percentage of their maximum acreage to plant closer toward 100% of their maximum acreage. Yet, ethanol demand does not explain why Corn Belt counties at very low percentages of their maximum acreage dramatically increase their acreage.

Table 15. Regression Results, Quantile with Latitudinal Expansion, IHS Variable Transformation, and State-Level Fixed Effects, M5

Variable	*q = 1/6*	*q = 1/3*	*q = 1/2*	*q = 2/3*	*q = 5/6*
IHS(LAGCORN)	0.955***	0.916***	0.889***	0.854***	0.785***
	(0.005)	(0.004)	(0.004)	(0.004)	(0.005)
IHS(LAGSOY)	0.012***	0.012***	0.012***	0.013***	0.010***
	(0.001)	(0.001)	(0.001)	(0.009)	(0.001)
IHS(CVCORN)	-0.091***	-0.083***	-0.057***	-0.040***	-0.041***
	(0.015)	(0.010)	(0.011)	(0.010)	(0.010)
IHS(CVSOY)	0.054***	0.033**	0.026**	0.039***	0.047***
	(0.018)	(0.012)	(0.012)	(0.012)	(0.011)
IHS(PXRATIO)	-0.499***	-0.461***	-0.398***	-0.267***	-0.238***
	(0.017)	(0.014)	(0.015)	(0.014)	(0.012)
IHS(PCRATIO)	0.148***	0.114***	0.087***	0.093***	0.142***
	(0.011)	(0.008)	(0.009)	(0.007)	(0.007)
IHS(ETHANOL)	-0.041***	-0.015***	0.017***	0.052***	0.054***
	(0.005)	(0.004)	(0.003)	(0.003)	(0.004)
IHS(ELR)	-0.016***	-0.008***	-0.004**	0.003*	0.015***
	(0.002)	(0.001)	(0.001)	(0.001)	(0.001)
IHS(ETEMP)	19.312***	11.620***	6.097***	0.468	-4.782***
	(1.510)	(1.021)	(0.995)	(0.991)	(0.956)
*IHS(ETEMP)*IHS(LAT)*	-7.259***	-5.056***	-3.272***	-1.903***	-1.450***
	(0.432)	(0.282)	(0.325)	(0.295)	(0.267)
*IHS(ETEMP)*IHS(ETEMP)*	-6.599***	-4.369***	-2.618***	-1.124***	-0.372
	(0.389)	(0.262)	(0.287)	(0.269)	(0.248)
*IHS(ETEMP)**	1.933***	1.337***	0.858***	0.491***	0.365***
*IHS(ETEMP)*IHS(LAT)*	(0.114)	(0.075)	(0.088)	(0.078)	(0.070)
IHS(EPREC)	1.071***	0.501*	0.280	-0.051	-1.130***
	(0.369)	(0.263)	(0.254)	(0.232)	(0.195)
*IHS(ETEMP)*IHS(EPREC)*	-0.304***	-0.157**	-0.097	-0.010	0.274***
	(0.097)	(0.069)	(0.067)	(0.062)	(0.052)

*Note: Fixed Effects and Time Dummies are omitted from reported results. Standard Errors in parentheses. *, **, and *** indicates significance at α = 0.01, 0.05, and 0.10 respectively.*

The sign of the parameter estimate for *ELR* changes from negative to positive across each quantile. This indicates that increases in expected loss ratios net of subsidies is a driver of increased acreage for counties already planting a high percentage of their maximum acreage, but increases in *ELR* do not result in increases to planted acreage when a county is planting a low percentage of their maximum acreage. Finally, the relationship between *ETEMP*, *EPREC*, and *ACRESCORN* vary by quantile. This is likely due to the geospatial pattern in values of *ACRESCORN* over time.

Re-Estimating the Complete Model for Fringe States

The complete model was re-estimated for the fringe states of Colorado, Kentucky, Mississippi, North Carolina, New York, Pennsylvania, Tennessee, and Texas. A comparison of regression results for the primary sample and the fringe producing regions is presented in Table 16.

LAGCORN is fairly robust across each sample. The parameter estimate recovered for *LAGSOY* is more than two times greater for counties in fringe states than Corn Belt counties. Yet, *CVSOY* is smaller and no longer statistically significant for counties in fringe states. Both relative prices and profitability have a greater impact on indexed acreage in counties in fringe states compared to Corn Belt counties. *ETHANOL* is not statistically significant for either sample, but the sign changes to positive for fringe counties. Notably, the parameter estimate for *ELR* is negative for counties in fringe states, which is contrary to the perception that counties in fringe states have historically exploited the RMA rating methodology and program mispricing.

The effect of changes in *ETEMP* and *EPREC* on *ACRESCORN* also varies by sample space. In Table 17 we present the change in *ACRESCORN* given a change in *ETEMP* from 21 to 22 degrees Celsius (the Corn Belt sample mean), and from 24 to 25 degrees Celsius (the fringe sample mean). These changes are calculated using the parameter estimates for *ETEMP* and *EPREC* recovered from the regression for each sample and are evaluated at the mean level of precipitation for each sample. Increases in expected temperature result in larger changes to *ACRESCORN* for all latitudes and levels of *ETEMP* in the Corn Belt than in the fringe states.

Table 16. Regression Results, Spatial Lag with Latitudinal Expansion, IHS Variable Transformation, and County-Level Fixed Effects, Corn Belt vs. Fringe States, M5

Variable	*Corn Belt*	*Fringe*
IHS(LAGCORN)	0.504***	0.604***
	(0.004)	(0.013)
IHS(LAGSOY)	0.008***	0.020***
	(0.002)	(0.004)
IHS(CVCORN)	-0.071***	-0.104***
	(0.012)	(0.033)
IHS(CVSOY)	0.047***	0.018
	(0.014)	(0.036)
IHS(PXRATIO)	-0.162***	-0.604***
	(0.014)	(0.051)
IHS(PCRATIO)	0.011	0.036
	(0.007)	(0.026)
IHS(ETHANOL)	-0.000	0.015
	(0.004)	(0.014)
IHS(ELR)	0.007***	-0.009*
	(0.002)	(0.005)
IHS(ETEMP)	28.867	-228.341***
	(20.743)	(78.534)
*IHS(ETEMP)*IHS(LAT)*	-7.170	54.628***
	(4.479)	(17.070)
*IHS(ETEMP)*IHS(ETEMP)*	-5.556**	28.627***
	(2.564)	(9.552)
IHS(ETEMP) IHS(ETEMP)*IHS(LAT)*	1.377**	-6.674***
	(0.550)	(2.138)
IHS(EPREC)	0.693**	3.685***
	(0.347)	(1.162)
*IHS(ETEMP)*IHS(EPREC)*	-0.209**	-0.996***
	(0.091)	(0.298)
n	19024	4756
σ^2	0.009	0.029
loglik	17087	1635
adj. r^2	0.926	0.903
ρ	0.557***	0.389***

*Note: Fixed Effects and Time Dummies are omitted from reported results. Standard Errors in parentheses. *, **, and *** indicates significance at α = 0.01, 0.05, and 0.10 respectively.*

Table 17. Change in ACRESCORN given changes in Temperature and Latitude, Corn Belt vs. Fringe

	35 North		*40 North*		*45 North*	
	Corn Belt	*Fringe*	*Corn Belt*	*Fringe*	*Corn Belt*	*Fringe*
Δ *ACRESCORN, T= 22 to 23*	-0.031	-0.023	-0.068	-0.0002	0.007	0.020
Δ *ACRESCORN, T= 24 to 25*	-0.026	-0.019	0.092	-0.004	0.011	0.009

Scenario Analysis

In Table 18 we present a scenario analysis for the Corn Belt sample using parameter estimates from the complete model. The change in *ACRESCORN* is reported for a 2%, 5%, 10%, 15%, and 25% increase in the mean value of each covariate, *ceteris paribus*. At the mean values of each covariate, indexed acreage is inelastic to changes in the covariate. It is important to note that while acreage response to our covariates is inelastic in the short run, yet in the long run the response is likely much more elastic.

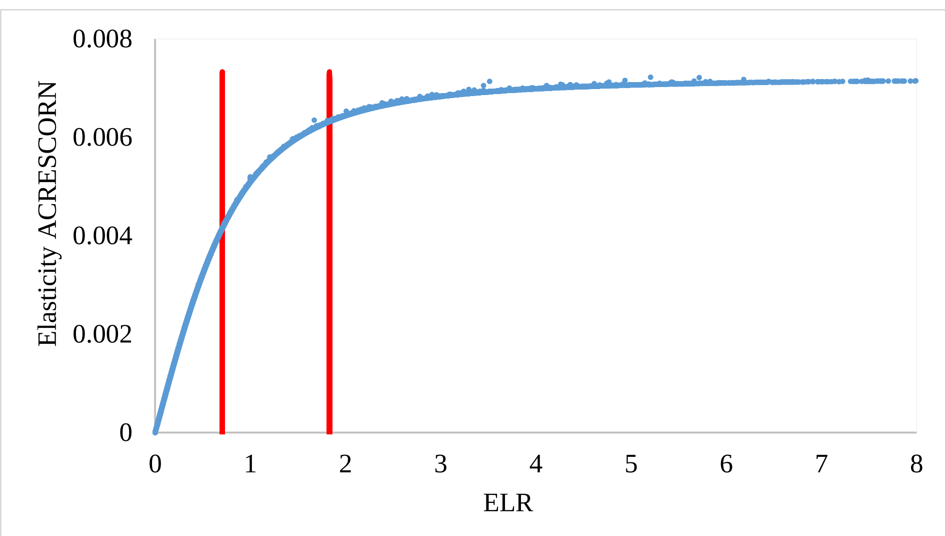
Our findings are consistent with the previous literature that finds inelastic acreage response due to changes in subsidies. Under an extreme case of a complete elimination of subsidies and a movement from the sample mean *ELR* of 1.442 to an actuarially fair rate of 1, a 30% reduction, the impact to indexed acreage would only be -0.004, or 0.4% reduction from the maximum acreage in the average county, *ceteris paribus*.

Table 18. Scenario Analysis, Spatial Lag with Latitudinal Expansion, IHS Variable Transformation, and County-Level Fixed Effects, M5

			Δ *ACRESCORN*				
Variable	$\hat{\beta}$	*Mean(Variable)*	*+2%*	*+5%*	*+10%*	*+15%*	*+25%*
LAGCORN	0.504	73.843	0.0220	0.0560	0.1080	0.1590	0.2530
LAGSOY	0.008	71.994	0.0004	0.0009	0.0017	0.0025	0.0040
CVCORN	-0.071	0.179	-0.0006	-0.0014	-0.0028	-0.0042	-0.0070
CVSOY	0.047	0.153	0.0003	0.0008	0.0016	0.0024	0.0039
PXRATIO	-0.162	2.357	-0.0067	-0.0165	-0.0323	-0.0475	-0.0763
PCRATIO	0.011	1.133	0.0004	0.0009	0.0018	0.0027	0.0043
ETHANOL	-0.0004	35.724	-0.0000	-0.0000	-0.0001	-0.0001	-0.0002
ELR	0.007	1.442	0.0002	0.0006	0.0013	0.0019	0.0030

In Figure 11 acreage elasticities are graphed across a range of observed values of *ELR*. The IQR for observed values of *ELR* is denoted by solid red lines. The higher the value of *ELR*, the more elastic the change to *ACRESCORN*. Yet overall, even at very high levels of *ELR*, the acreage response is inelastic.

Figure 11. ELR vs. Elasticity of ACRESCORN



In Figure 12 acreage elasticities are graphed across observed values of *ETEMP*. Each elasticity is evaluated at the corresponding observed level of *EPREC* and *LAT*, and the fitted value of *ACRESCORN* recovered from the model. Overall, acreage response to temperature is inelastic. As expected, for higher values of *ETEMP*, elasticities are negative indicating further increases in mean summertime temperatures would result in a reduction to acreage. This result has acreage implications when considering future climate change forecasts for the warmest regions of the sample, which we discuss below. Surprisingly, for the lowest values of *ETEMP*, acreage elasticity is also negative. For values of *ETEMP* that wall within the IQR range, the sign of acreage elasticity varies from positive to negative, although overall inelastic, indicating corresponding values of *EPREC* and *LAT* must be considered when determining acreage response to temperature.

Similar to *ETEMP*, acreage response to changes in *EPREC* are inelastic across observed values of *EPREC*. Unlike *ETEMP*, acreage response to changes in *EPREC* is negative across observed values, which is a surprising result, given our hypothesis that farmers consider drought risk when making acreage decisions. It could be the case that using the parameter estimate recovered for the effect of *EPREC* on mean values of *ACRESCORN* does not result in fully informative acreage elasticity estimates across all values of *EPREC*.

Figure 12. ETEMP vs. Elasticity of ACRESCORN

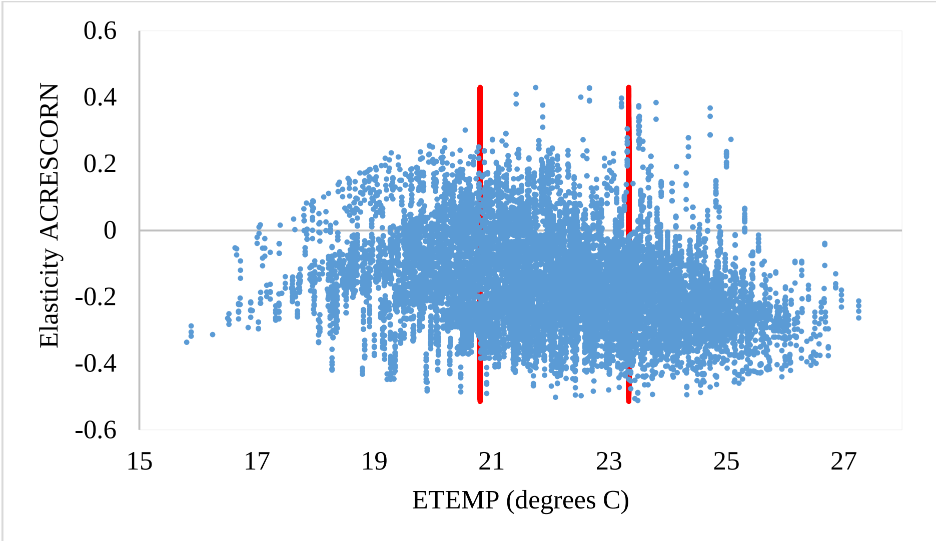
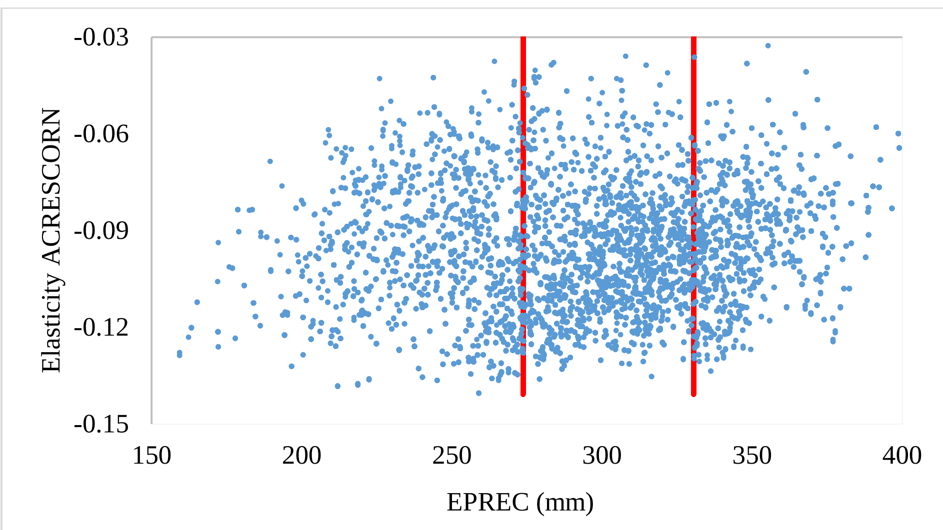


Figure 13. EPREC vs. Elasticity of ACRESCORN

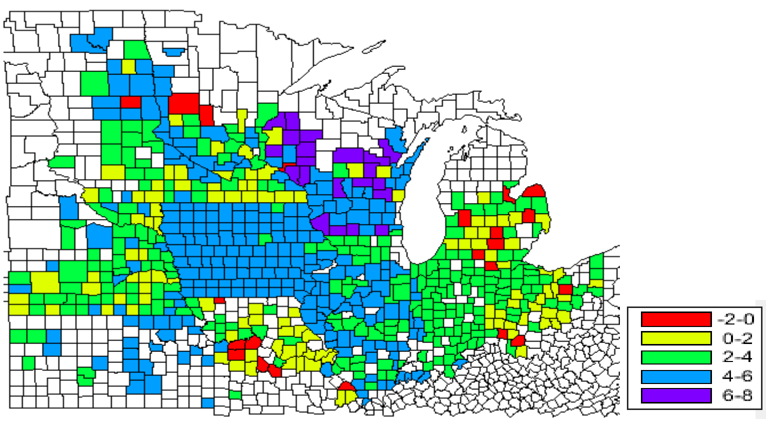


Climate Change Scenario Analysis

Some economists and policy makers perceive climate change as a *future* risk to agriculture, yet the past fifty years in the US Midwest has been characterized by warming. The difference between historical climate change and predicted climate change is that it is generally acknowledged historical warming was beneficial to agriculture. Furthermore, the past fifty years have been characterized by tremendous technological progress. The concern for agricultural production in the future is that technological progress will plateau and/or climate change will outpace continued technological progress. Given the interest in agricultural production in a future climate, an acreage scenario analysis given predicted changes in summertime temperatures and precipitation in the Midwest is warranted.[10]

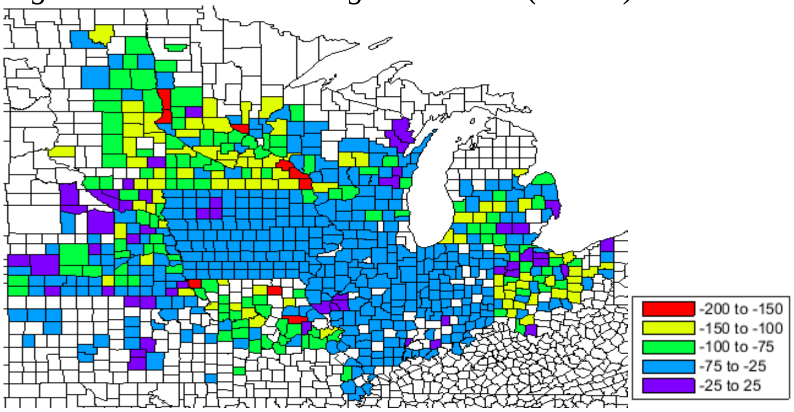
Figures 14 and 15 show predicted changes to *ETEMP* and *EPREC* between 2013 and 2050 for the 656 counties in our sample. *ETEMP* is predicted to increase in all but 21 of the 656 counties. *EPREC* is predicted to decrease in all but two counties. The average predicted change in *ETEMP* is 3.306 degrees Celsius, and the average predicted change to *EPREC* is -63.48mm.

Figure 14. Predicted Change in ETEMP (Δ degrees Celsius) between 2013 and 2050



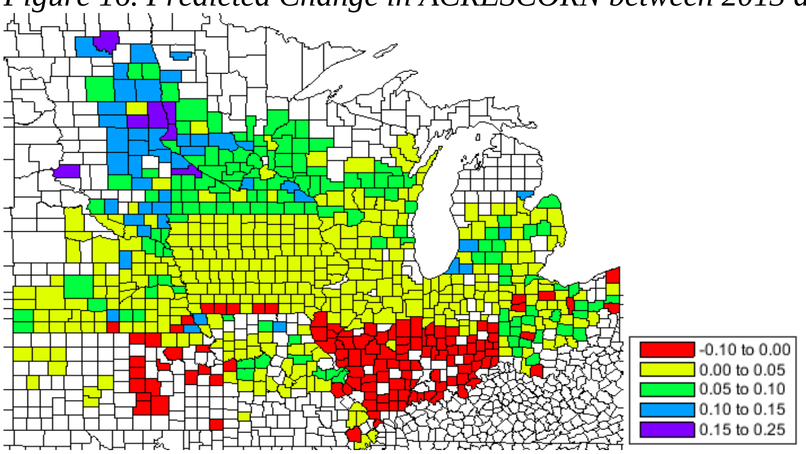
[10] Projections for *ETEMP* and *EPREC* are generated using data from ClimateWizard.org. Climate Wizard processes climate change projections generated by 16 different general circulation models (GCMs) for the 2050s and the 2080s. The sixteen general circulation models predicting mean changes in temperature and precipitation for 2040 - 2069 and 2070 - 2099 as compared to a 1961 – 1990 base period are: BCCR-BCM2.0, CGCM3.1 (T47), CNRM-CM3, CSIRO-Mk3.0, GFDL-CM2.0, GFDL-CM2.1, GISS-ER, INM-CM3.0, IPSL-CM4, MIROC3.2 (medres), ECHO-G, ECHAM5/ MPI-OM, MRI-CGCM2.3.2, CCSM3, PCM, and UKMO-HadCM3. *ETEMP* and *EPREC* is constructed by using predicted changes to temperature and precipitation at the 50th percentile consensus among these 16 models at the medium emission scenario.

Figure 15. Predicted Change in EPREC (Δ mm) between 2013 and 2050



In scenario 1 (Figure 16) we map the predicted change in *ACRESCORN* between 2013 and 2050. We use forecasted values for *ETEMP* and *EPREC* and hold all other covariates at their 2013 levels. The model predicts marginal changes to *ACRESCORN* between -0.10 and 0.25. The largest increases to acreage occur in the Northwestern portion of the sample space, whereas the decreases in acreage occur in the Southeastern portion of the sample space. This is consistent with the assumption that cooler producing regions may be able to better withstand future warming.

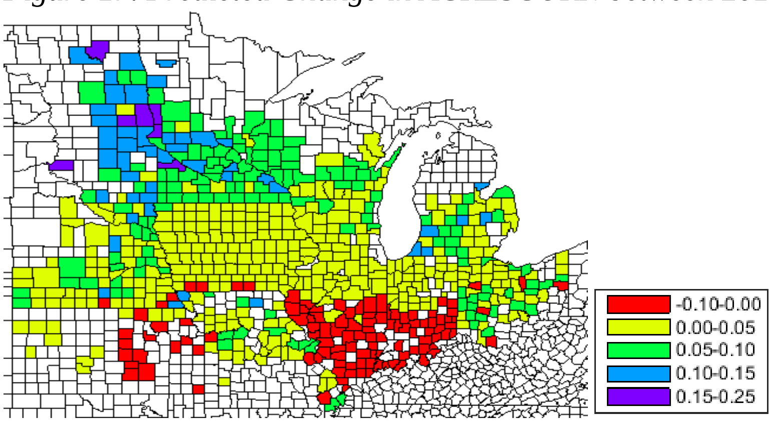
Figure 16. Predicted Change in ACRESCORN between 2013 and 2050, Scenario 1



Given the interest in forecasting yields and yield risk under future climate change, in scenario 2 (Figure 17) we map the change in *ACRESCORN* between 2013 and 2050 predicted by our model from using forecasted values for *ETEMP* and *EPREC*, and we increase yield risk by

25% for every county in the sample. As corn and soybeans are substitute commodities we assume technological improvements to offset any negative effects of climate change are the same for both crops. Because the percentage increase in yield risk is the same for both corn and soybeans the relative price ratio of soybeans to corn remains the same as 2013. In spite of shocking yield risk for both commodities by 25%, the acreage effect from temperature and precipitation outweighs the effects of increased yield risk, and results remain nearly identical to scenario 1.

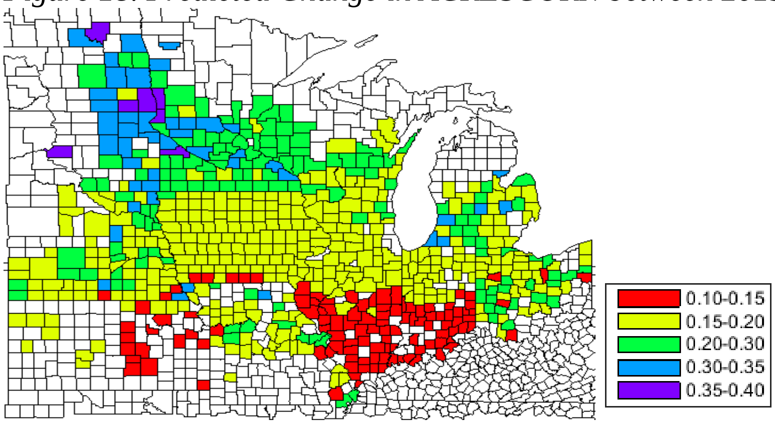
Figure 17. Predicted Change in ACRESCORN between 2013 and 2050, Scenario 2



Lastly, in scenario 3 (Figure 18) we map the change in *ACRESCORN* between 2013 and 2050 given forecasted changes in *ETEMP* and *EPREC*. In this scenario we assume that both corn and soybean yield risk have increased, but that the increase in corn yield risk exceeds the increase in soybean yield risk. For example, we increase corn yield risk by 25% for all counties, but only increase soybean yield risk by 10%. Furthermore, due to corn yield risk becoming relatively high in comparison to soybeans we assume corn producers are compensated by the market for this extra risk. We assume the relative price ratio of soybeans to corn decreases by 50% between 2013 and 2050, from 2.284 to 1.142. Although the pattern of changes in acreage remains identical to scenarios 1 and 2, the effect of changes in temperature and precipitation, increased yield risk, and decreased relative prices is an *increase* in corn acreage for all counties in the sample. Hence relative price is a driver of acreage decisions in comparison to expected weather and yield risk.

Across all scenarios, dramatic changes to temperature, precipitation, yield risk, and relative prices result in marginal changes to acreage. Parameter estimates recovered from the model are more reflective of short run elasticities, so it is possible that over a longer time series such as 2013 to 2050 the overall acreage effect would be much greater.

Figure 18. Predicted Change in ACRESCORN between 2013 and 2050, Scenario 3



CHAPTER VI

CONCLUSION

We constructed a comprehensive model for planted corn acreage in the US Midwest that accounted for the impacts of crop insurance subsidization and expected weather on farmer planting behavior, contributing to two important policy themes in the current agricultural economic literature. Indexed corn acreage planted was regressed on lagged corn and soybean acreage, corn and soybean yield risk, the relative price ratio of soybeans to corn, a producer price to cost ratio, producer subsidy take, lagged ethanol consumption, expected loss ratios, expected subsidy take, expected growing season maximum temperature, and expected growing season precipitation. The primary sample area included the twelve "Corn Belt" states, and the time period used for the analysis was 1985 to 2013 inclusive. This analysis focused on the influence of FCIP subsidies and perceived program mispricing on corn acreage *planted* in the Midwest, rather than on acreage *enrolled* in the FCIP. Unlike the previous literature which accounted for weather in models of planted acreage, we constructed *expected* values of temperature and precipitation to incorporate into our model. These variables better reflected the rational expectation of a farmer in comparison to observed weather that is subject to extreme intra and inter annual variability.

A spatial lag specification with a latitudinal expansion of *ETEMP* and $ETEMP^2$ and an inverse hyperbolic sine transformation of both the dependent and explanatory variables was selected as the complete model for *ACRESCORN*. The combination of covariates that performed best both in and out of sample was M5 (*LAGCORN*, *LAGSOY*, *CVCORN*, *CVSOY*, *PXRATIO*, *PCRATIO*, *ELR*, *ETEMP*, *ETEMP*LAT*, $ETEMP^2$, $ETEMP^2$**LAT*, *EPREC*, *ETEMP*EPREC*, *ETHANOL*, *TIME*).

LAGCORN was positive and significant, indicating trends within the data where county/years on average exhibit periods of increasing or decreasing planting. Perhaps this is due in part by the fixity of farming and/or the limited uses for the land. *LAGSOY* was positive and significant, and may be indicative of crop rotation. *CVCORN* was negative and significant, indicating farmers plant less corn when yield risk increases. This is likely because there are substitute commodities available, namely soybeans. Similarly *CVSOY* was positive and significant indicating if soybean production risk is increased farmers may abandon crop rotation practices and plant more corn. *PXRATIO* was negative and significant, indicating if soybeans become relatively rich to corn farmers will abandon crop rotation practices in the short-run. *PCRATIO* was positive and significant, meaning for an increase in farm profitability farmers will plant more acres of corn. Perhaps this is because land is otherwise left fallow for environmental purposes or there are other better uses for farm capital, labor, or time.

Models using *ELR* (as a proxy for both FCIP subsidization and perceived program value) were preferred to models that used *ESUBTAKE* and *EGLR* since these models performed better in out of sample validation. We find that increases in *ELR* result in marginal increases to *ACRESCORN*. We conclude that farmers' expectations of growing season weather (influenced by historical climate change) are significant and overlooked factors driving changes in farmer planting behavior, particularly in the more northern and fringe producing regions for US corn.

A graph of acreage elasticities shows the most elastic response by farmers to increases in *ELR* is in counties where *ELR* is highest, yet overall *ACRESCORN* is very inelastic to changes in *ELR*. Consistent with the previous literature, our results suggest that an aggressive re-pricing of the FCIP or reductions in program subsidies would only result in a marginal impact to planted acreage. At an extreme, if corn subsidies were eliminated entirely and the program was priced to be actuarially fair, our results suggest the average decrease to *ACRESCORN* in each county would only be 0.4%.

The results of a quantile regression suggest there is a statistically significant difference in acreage response to subsidy adjusted loss ratios according to the quantile of our dependent variable. However, at odds with the existing literature and our own intuition, we find that in county/years of low observed values of *ACRESCORN* the sign of the parameter estimates for *ELR* are *negative*. It is in county/years in the highest quantile of our dependent variable where farmers are most responsive to increases in subsidy take or perceived program mispricing. That is, increases in *ELR* or *ESUBTAKE* cannot be attributed to conversion of environmentally sensitive or fringe production. In fact, the results from the quantile regression suggest it may real and relative prices that are driving the conversion of land.

In future development of this research there are several data issues to address. First, the study would benefit from a better urbanization measure, since state-level urbanization percentages from the US census is too spatially aggregated to capture the urbanization effects we know to be present in our sample area, in particular in Illinois. Furthermore, as US census data is only calculated every ten years, we linearly interpolated between census years to proxy for urbanization trends. It would be ideal to get urbanization data at a more spatially and temporally refined scale to separate the effects of urbanization from our county-level fixed effect terms. Second, the analysis would benefit from an improved price to cost index. The ERS Midwest annual price to cost index level was imposed on every county in our sample, but an index that varies spatially as well as temporally would be ideal. We will look to construct our own index from NASS and ERS farm financial data. In future research we plan to explore the soil expansion model in more depth. Instead of expanding by NCCPI soil quality index, we will try using more explicit measures of soil quality such as soil organic carbon matter or moisture content at certain depths.

The statistical significance of our crude expected weather term warrants further consideration into how farmers shape their expectations of weather, as the results of our econometric models are only as good as our assumptions regarding farmers' expectations. Being a land grant University we are uniquely suited to conduct a survey to better identify how farmers shape their expectations for upcoming growing season weather. The practice in econometric literature for other expectation variables, such as price, is to weight previous experiences. However, this may not work well for weather data since there is so much variance in the observed data; that is, model results in outlier years might not be a true reflection of a farmer's expectations, as the profit-maximizing farmer is likely to consider the year-over-year variability that is common in weather. Conducting a farmer-survey to better understand how farmers shape their expectations may be warranted. Questions that should be asked include the following: do farmers consider wintertime temperature or precipitation when forming an expectation as to summertime weather? How many years of historically observed weather does a farmer consider when shaping their expectations for the upcoming growing season? What seasonal weather forecasts (if any) do farmers deem most reliable? A quick review of the literature suggests no survey of this kind has been conducted. This survey would make for a nice behavioral piece for agricultural economists, and the information could serve as a basis to construct a farmer growing season forecast index from historical data.

Since structural changes such as increased subsidization to the FCIP are often contentious among economists, there are many other dimensions of the program to analyze in future research. The results of this study suggest that the FCIP has not been an economically significant factor in the conversion of environmentally sensitive and fringe land to planted corn. A research topic that is related to this research is an analysis of the influence of the FCIP or other risk management programs on farmers' willingness and/or success in adapting to climate change. Although adaptation is a crowded space in the literature, we could model historical yield risk (coefficient of variation) for a crop in two countries; one which has a robust insurance market for farmers, such as the US, and a second country which lacks an insurance market, such as Australia. Yield risk could be regressed on observed weather and other factors such as soil quality, irrigation, and fertilizer use to see which country can attribute more yield risk to extreme weather.

Another extension of this analysis would be to model how a county goes from 0 to 100% planted once acreage allocations have been decided. NASS publishes a crop progress report and reports over the course of 7 or 8 weeks how quickly a county moves from 0 to 100% planted. We plan on using a recursive model where each week's planting is a function of the previous week's planting, the previous week's weather, and the farmer's perception of weather in the upcoming week, with planting subject to labor, equipment, and time constraints.

This analysis contributes to the previous literature on planted acreage in several ways. First we contribute with improved econometric specifications. By utilizing spatial methods we best control for spatial dependency in the data. To our knowledge there is no previous literature utilizing quantile regressions to analyze FCIP and planted acreage. Our results demonstrate that acreage response to changes in the FCIP are not robust across different quantiles of our dependent variable; in fact, the results of the quantile regression completely refute the hypothesis that fringe-producing regions are most responsive to changes in subsidies or expected returns.

We provide more consideration to the dependent variable and explanatory variables than previous studies of acreage. By indexing acreage we allow more flexibility in terms of the impact county land side has on planted acreage, unlike other studies which either regress planted acreage on county land acres or try to capture the effect of land size in the fixed effect term. Compared to previous studies which have analyzed the effect of the FCIP on planted acreage, we gave more consideration to the individual components of the FCIP that impact acreage, thereby providing more insightful conclusions for policy makers. Furthermore, we account for the impact of expected weather on planted acreage. The methods utilized in this study could be extended to other analyses of planted acreage or production in which spatial dependencies exist.

REFERENCES

Anselin, Luc. "Spatial effects in econometric practice in environmental and resource economics." *American Journal of Agricultural Economics* (2001): 705-710.

Anselin, Luc. *Spatial Econometrics: Methods and Models*. Vol. 4. Springer, 1988.

Chen, Xiaoguang, and Hayri Önal. "Modeling agricultural supply response using mathematical programming and crop mixes." *American Journal of Agricultural Economics* 94.3 (2012): 674-686.

Claassen, Roger, Joseph C. Cooper, and Fernando Carriazo. "Crop Insurance, Disaster Payments, and Land Use Change: The Effect of Sodsaver on Incentives for Grassland Conversion." *Journal of Agricultural and Applied Economics* 43.2 (2011): 195.

Glauber, Joseph W. "The Growth of The Federal Crop Insurance Program, 1990–2011." *American Journal of Agricultural Economics*. 95.2 (2013): 482–488.

Goodwin, Barry K., and Vincent H. Smith. "An ex post evaluation of the conservation reserve, federal crop insurance, and other government programs: program participation and soil erosion." *Journal of Agricultural and Resource Economics* (2003): 201-216.

Goodwin, Barry K., Monte L. Vandeveer, and John L. Deal. "An empirical analysis of acreage effects of participation in the federal crop insurance program." *American Journal of Agricultural Economics* 86.4 (2004): 1058-1077.

Gregory, Peter J., John SI Ingram, and Michael Brklacich. "Climate change and food security." *Philosophical Transactions of the Royal Society B: Biological Sciences* 360.1463 (2005): 2139-2148.

Huang, Haixiao, and Madhu Khanna. "An econometric analysis of US crop yield and cropland acreage: implications for the impact of climate change." *Available at SSRN 1700707* (2010).

Just, Richard E. "An investigation of the importance of risk in farmers' decisions." *American Journal of Agricultural Economics* 56.1 (1974): 14-25.

Kaiser, Harry M., Susan J. Riha, Daniel S. Wilks, David G. Rossiter, and Radha Sampath. "A farm-level analysis of economic and agronomic impacts of gradual climate warming." *American Journal of Agricultural Economics* 75.2 (1993): 387-398.

Koenker, Roger, and Gilbert Bassett Jr. "Regression quantiles." *Econometrica: Journal of the Econometric Society* (1978): 33-50.

Kramer, Randall A. "Federal crop insurance 1938-1982." *Agricultural History* (1983): 181-200.

LeSage, James, and Robert Kelley Pace. *Introduction to Spatial Econometrics*. CRC press, 2010.

Lobell, David B., Wolfram Schlenker, and Justin Costa-Roberts. "Climate trends and global crop production since 1980." *Science* 333.6042 (2011): 616-620.

Lubowski, Ruben N., Bucholtz, S., Claassen, R., Roberts, M. J., Cooper, J. C., Gueorguieva, A., & Johansson, R. (2006). *Environmental effects of agricultural land-use change*. Washington, DC, USA: US Department of Agriculture, Economic Research Service.

Lubowski, Ruben N., Andrew J. Plantinga, and Robert N. Stavins. "What drives land-use change in the United States? A national analysis of landowner decisions." *Land Economics* 84.4 (2008): 529-550.

Miao, Ruiqing, Hongli Feng, and David A. Hennessy. "Land Use Consequences of Crop Insurance Subsidies." *Agricultural & Applied Economics Association's 2011 AAEA & NAREA Joint Annual Meeting, Pittsburgh, Pennsylvania, July*. 2011.

Miranda, Mario J., and Joseph W. Glauber. "Systemic risk, reinsurance, and the failure of crop insurance markets." *American Journal of Agricultural Economics* 79.1 (1997): 206-215.

Muth, John F. "Rational expectations and the theory of price movements."*Econometrica: Journal of the Econometric Society* (1961): 315-335.

Nerlove, Marc. "Estimates of the elasticities of supply of selected agricultural commodities." *Journal of Farm Economics* 38.2 (1956): 496-509.

Nowshirvani, Vahid F. "A modified adaptive expectations model." *American Journal of Agricultural Economics* 53.1 (1971): 116-119.

Schlenker, Wolfram, and Michael J. Roberts. "Nonlinear temperature effects indicate severe damages to US crop yields under climate change." *Proceedings of the National Academy of Sciences* 106.37 (2009): 15594-15598.

Skees, Jerry R., and Barry J. Barnett. "Conceptual and practical considerations for sharing catastrophic/systemic risks." *Review of Agricultural Economics* 21.2 (1999): 424-441.

Tegene, Abebayehu, Wallace E. Huffman, and John A. Miranowski. "Dynamic corn supply functions: a model with explicit optimization." *American Journal of Agricultural Economics* 70.1 (1988): 103-111.

Vesterby, Marlow, Shawn Bucholtz, Alba Baez, and Michael J. Roberts. “Major uses of land in the United States, 2002”. *USDA-ERS* (2006).

Walters, Cory G., C. Richard Shumway, Hayley H. Chouinard, and Philip R. Wandschneider. "Crop insurance, land allocation, and the environment."*Journal of Agricultural and Resource Economics* 37.2 (2012): 301.

Westcott, Paul C., and C. Edwin Young. "Farm program effects on agricultural production: coupled and decoupled programs." *Burfisher, M. and Hopkins, J., Editors* (2004).

Woodard, Joshua D. “AgDB Data Warehouse”, Cornell University, 2014, Available online at: https://agfinance.dyson.cornell.edu/AgRiskManagement/ResAgDataQuery

Woodard, Joshua D., Leslie J. Verteramo Chiu, and Alyssa P. Miller. "Impacts of Technological Assumptions on Agricultural Yield Forecasts under Climate Change." Working paper, Cornell University, 2014.

Woodard, Joshua D., and Philip Garcia. "Weather derivatives, spatial aggregation, and systemic risk: implications for reinsurance hedging." *Journal of Agricultural and Resource Economics* (2008): 34-51.

Wright, Brian Davern. "Multiple Peril Crop Insurance." *Choices* 29.3 (2014).

Wu, JunJie. "Crop insurance, acreage decisions, and nonpoint-source pollution." *American Journal of Agricultural Economics* 81.2 (1999): 305-320.

Wu, JunJie, and Richard M. Adams. "Production risk, acreage decisions and implications for revenue insurance programs." *Canadian Journal of Agricultural Economics/Revue Canadienne d'Agroeconomie* 49.1 (2001): 19-35.

Young, C. Edwin, Monte L. Vandeveer, and Randall D. Schnepf. "Production and price impacts of US crop insurance programs." *American Journal of Agricultural Economics* 83.5 (2001): 1196-1203.

Young, C. Edwin, and Paul C. Westcott. "How decoupled is US agricultural support for major crops?." *American Journal of Agricultural Economics* 82.3 (2000): 762-767.

Zacharias, Thomas P., and Keith J. Collins. "Ten Considerations Regarding the Role of Crop Insurance in the Agricultural Safety Net." *Choices* 28.3 (2013).